Reciprocity, Incomplete Information and Relational Contracts

Dissertation
submitted to the Faculty of Economics,
Business Administration and Information Technology
of the University of Zurich

to obtain the degree of
Doktor der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by

André Volk
from Germany

approved in February 2015 at the request of

Prof. Dr. Armin Schmutzler
Prof. Dr. Björn Bartling

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 11.02.2015

Chairman of the Doctoral Board: Prof. Dr. Josef Zweimüller

to my parents

Preface

My interest in human behaviour has sparked and fueled this writing. I find that economics, enriched with insight from psychology, can be very helpful in organizing individual and interactive behaviour. I especially value the attempt to conduct a line of argument transparently, by explicitly stating the assumptions underlying a theory and by deriving the conclusions using the rules of logic. I consider the particular assumptions, for instance with regard to the rationality and preferences of individuals, as well as the use of formal language as characterizing a discipline which provides a way of thinking; a window into the world, much like the view through a lens which renders certain structures sharp while others remain blurry or invisible. Thinking about the economic way of thinking has informed my understanding of its power and range. This has shaped my ideas and the research presented in this thesis.

I wish to thank the numerous people who supported me and my work throughout the time of my dissertation. First and foremost, I feel grateful to Armin Schmutzler for his very generous support, his encouragement and patience, and the many exchanges of ideas about my research. I am indebted to Nick Netzer for teaching me how to conduct theoretical research and to think like a theorist. I am thankful to Holger Herz for showing me how to run experiments and for his advice on my experimental research. My ideas have also benefited from the many interesting conversations with current and former members at the Department of Economics at the University of Zurich, especially Björn Bartling, Charles Efferson, Ernst Fehr, Andreas Hefti, Alexey Kushnir, Michel Maréchal, Konrad Mierendorff and Tony Williams.

I am deeply grateful to my friends and family whose loving care and support has helped me reach my goals with regard to this thesis and develop myself as a person.

André Volk
Zürich, September 2014

Contents

Introduction

Economic research often applies the assumption that individuals are rational and value only their own material benefit. This view has facilitated many powerful insights in economics. Unlike the model of material self-interest, the theory of intention-based preferences for reciprocity, as first introduced by Rabin (1993), presumes that individuals are willing to sacrifice own material benefit in order to be kind to those who are kind to them and to be unkind to those who are unkind to them. This behavioural pattern has received strong empirical support.[1] The model of intention-based preferences also differs conceptually from the theory of material self-interest. Preferences are exclusively defined over final outcomes in the latter model. By contrast, kindness sensations in the intention-based model are determined within the underlying game. Hence, the evaluation of outcomes by individuals who care about intentions can depend on the structure of the game in which they participate. This procedural aspect of intention-based preferences has interesting implications for the study of mechanism design, which addresses how games can be designed to achieve desired allocations.[2]

Bierbrauer and Netzer (2014) first studied intention-based preferences for reciprocity in a mechanism design framework. The authors provide conditions such that the implementation under the assumption of material self-interest is robust to a broad class of other-regarding preferences, including the intention-based model. Moreover, they illustrate how the procedural properties of intentions-based social preferences can be exploited for the design of mechanisms. In particular, the authors show how kindness sensations can be affected by the construction of the mechanism such that individuals behave in the way desired by the designer. As a result, they provide conditions such that any Pareto efficient social choice function can be implemented with voluntary participation.

Chapter 1 of this thesis is the result of joint work with Nick Netzer. We build on the framework developed by Bierbrauer and Netzer (2014) and study two robustness concerns for the implementation with intention-based preferences for reciprocity. First, we provide a notion of robustness concerning the formation of beliefs about others' intentions under incomplete information. Each player in our model is privately informed about his material payoff type, analogous to the conventional framework with materially self-interested individuals. This implies that a player's intended kindness depends on his privately observed type. One player therefore cannot pin down another player's true kindness as he is uncertain about this player's actual type. Very little is known about how individuals

[1]See e.g. Sobel (2005) for a survey article on reciprocity. The empirical relevance of intentions has been documented by e.g. Andreoni et al. (2002) and Falk et al. (2003, 2008).

[2]See e.g. the discussion in Bierbrauer and Netzer (2014).

attribute intentions under asymmetric information and a designer may not want to rely on a specific and possibly misspecified model. We therefore propose a notion of implementation with the property of type-invariance, which requires the equilibrium kindness of each player to be independent of his type. This property renders each player's intentions transparent, despite the presence of asymmetric information.

Second, we propose a notion of ex-post implementation for the case of intention-based preferences for reciprocity. Our concept requires that a player would not want to deviate from his equilibrium behaviour, even if he were informed about all other players' private information. Having information about others' actual types would allow each player to infer their true intentions. Moreover, each player could learn his own ex-post material payoff, analogous to the case of ex-post implementation with materially self-interested individuals. Our definition of ex-post implementation implies that a player would not want to revise his interim decision if he were given such opportunity ex-post. Also, our concept rules out that a player anticipates at the interim stage that he will regret his decision when he learns all other players' private information ex-post.

As our main result, we show that implementation with type-invariance as well as ex-post implementation is possible for any materially Pareto efficient social choice function which provides insurance to the agents. In a direct mechanism, insurance is provided if unilateral deviations from truth-telling do not affect any other player's material payoff (Bierbrauer and Netzer, 2014). This insurance property provides the basis for the implementation with type-invariance as it induces zero kindness if players tell the truth in the direct mechanism. We exploit the procedural nature of intention-based preferences by adding actions to such direct mechanism in order to shape kindness sensations obtained from truth-telling while, at the same time, keeping them type-invariant. We show how mechanisms can be constructed in this way and equilibrium kindness can be tuned such that private and social interests become aligned at the interim and even at the ex-post stage. We further illustrate that our main result holds true even if participation to the implementing mechanism is voluntary.

Incomplete information is considered one of the main impediments to efficiency, not only within the mechanism design framework studied in Chapter 1, but also in economics more generally. Much of the research on incomplete information presumes that incentives are provided by the use of explicit, formal contracts. Many real world economic relationships, however, are governed by informal, implicit agreements (see e.g. MacLeod (2007) and Malcomson (2012b) for surveys). In contrast to formal contracts, such relational "contracts" cannot be enforced by neutral third parties. Yet, the literature suggests that they can be self-enforcing within an ongoing relationship: individuals follow an implicit agreement to cooperate if they believe that deviation to the own benefit will be punished in the course of future interaction. In this sense, cooperation and the provision of efficiency are incentivized by the shadow of the future.

In Chapter 2, which is result of joint work with Holger Herz and Armin Schmutzler, we study experimentally and theoretically how incomplete information affects the economic performance of relational contracts. Informational asymmetries pose a challenge for implicit agreements as they can preclude one player to tell whether another cooperates and honours a relational contract or takes an unfair advantage to the own benefit. Doubt about whether an agent plays fair may lead to allegations and possibly false accusations which can destabilize a relationship. This suggests that incomplete information adversely affects the functioning of relational contracts and leads to economically inferior outcomes.

Only little is known about the effect of asymmetric information on relational contracts and previous research has primarily taken a theoretical perspective. We contribute to the literature by providing empirical evidence complemented by a theoretical analysis. We study repeated principal-agent relationships in which the agent holds private information about his effort costs being either high or low. In the beginning of a relationship, the agent sends a signal to his principal concerning his costs which, however, need not be truthful. If a principal receives a high cost signal, then she cannot be sure whether her agent tells the truth or is assigned to low costs and pretends to have high costs. As a consequence, such principal cannot be sure whether her agent plays fair or exploits the informational asymmetry to the own advantage.

Our experimental evidence shows that a large fraction of low cost agents behaves dishonestly. Principals doubt the truthfulness of high cost signals. However, this does not lead to adverse consequences for efficiency compared to the case of high costs in our control treatment with complete information, where doubts are absent by construction. This is surprising in light of the fact that dishonesty causes potential for conflict in a relationship: dishonest low cost agents induce substantial payoff inequality and severely harm their principals' payoffs. Overall, we find that asymmetric information can but must not necessarily harm efficiency. We show in addition that the observed behaviour can be organized by a theoretical model of relational contracts.

Chapter 3 of this thesis further contributes to the empirical understanding of relational contracts. It begins with the observation that many ongoing economic relationships are structured as a sequence of transactions, where each transaction by itself takes place sequentially. In a relationship between a principal and her agent, a transaction may consist of the principal providing a wage first and the agent exerting effort thereafter. Alternatively, the agent may act as the first- and the principal as the second-mover in a stage game within the repeated interaction. This chapter explores experimentally how the order of moves, in which a transaction takes place, can be utilized to enhance the functioning of relational contracts with regards to the provision of efficiency.

If relational contracts are driven by the shadow of the future, as primarily suggested by the literature, then the order of moves may not effect efficiency: if the relationship is ongoing, deviations from the relational contract can be punished irrespective of whether the deviating party acts as the first- or the second-mover. This suggests that incentives to conduct efficient transactions do not depend on who moves first. However, this hypothesis does not account for the possibility that individuals may take a narrow rather than a broad perspective on their decisions (see Read et al. (1999) for a survey). In particular, the parties may narrowly focus on the outcome of the current stage game rather than consider the outcome of the relationship as a whole. The evidence presented in Chapter 3 suggests that such narrow bracketing can lead to consequences of the order of moves on efficiency.

The experimental data shows that efficiency in stage games in which the agent acts as the first-mover is substantially higher than stage games in which the principal moves first. This result emerges if the order of moves persists and also if it alternates over the course of a relationship. The observed behaviour is consistent with the prediction of narrow bracketing: second-movers claim more than an equal share of the surplus in a stage game, despite the fact that the relationship is ongoing. This implies that principals who act in the position of the first-mover pay higher than fair wages, which substantially exceed the costs of effort provision. Hence, principals moving first have to make much larger investments than agents moving first in order to induce a transaction with given efficiency level. If first-movers hesitate to invest large amounts, this may explain why transactions in which the agent moves first generate higher efficiency compared to transactions in which the principal acts as the first-mover.

Chapter 1

Robustness and Ex-Post Implementation with Intention-Based Reciprocity

Intention-based reciprocity is an important motivation for human behavior, and it can be exploited in the design of economic allocation mechanisms. In this chapter, we address questions of robustness that arise in the context of asymmetric information about intentions. We propose allocation mechanisms that eliminate uncertainty about the players' intentions, by making all types of each player equally kind, and we investigate a first notion of ex-post fairness implementation, based on the property that learning about a player's type does not change the perception of that player's intention in such mechanisms. We show that efficient social choice functions which provide payoff insurance to the agents can be implemented in this way, with or without voluntary participation constraints.

1.1 Introduction

Private information about payoffs constitutes a key ingredient to the literature on mechanism design. We consider a mechanism design model where behavior is driven by material as well as psychological motives. In particular, we follow the framework of Bierbrauer and Netzer (2014) and equip agents with intention-based preferences for reciprocity (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004). Private information about material payoffs then gives rise to private information about intentions: players cannot fully determine the other players' intended kindness, as they are lacking information about these players' types.[1] We contribute to the existing literature by addressing two problems that may arise in this context.

First, our understanding of the attribution of intentions under uncertainty is still limited. We might, for instance, not be able to discard the hypothesis that opponent types with bad intentions are more salient than those with good intentions, and hence are overweighted in the process of belief formation.[2] Therefore, while we could assume that players treat intentions like material payoffs and form an expected value at the interim stage, here we take a broader approach. We propose an equilibrium notion that requires intentions to be type-invariant, and we construct mechanisms that implement certain materially efficient social choice functions in this way. An equilibrium where each type of a player displays the same kindness remains robust for a large class of assumptions about the attribution of intentions under asymmetric information.

Second, if players were informed about all private information after taking their actions, then they could infer the other players' true intentions and may regret their decisions, as they may have acted differently had they known these intentions beforehand. Such psychological regret can be a concern in the same way and in addition to material ex-post regret, linked to the uncertainty about material payoff types, as addressed in the theory of ex-post implementation.[3] We therefore propose a notion of ex-post fairness implementation which requires that all players would want to stick to their interim decisions even if they were informed about all private information ex-post. We again utilize the property of

[1] Models of intention-based preferences rely on the framework of psychological game theory (Geanakoplos et al., 1989; Battigalli and Dufwenberg, 2009). Most of the papers that develop or apply models where intentions matter do not explicitly consider asymmetric information (e.g. Rabin, 1993; Dufwenberg and Kirchsteiger, 2000, 2004; Charness and Rabin, 2002; Falk and Fischbacher, 2006; Cox et al., 2007; Segal and Sobel, 2007, 2008; Hahn, 2009; Sebald, 2010; Nishimura et al., 2011; Dufwenberg et al., 2011, 2013; Hoffmann and Kolmar, 2013; Netzer and Schmutzler, 2013). Bierbrauer and Netzer (2014) and von Siemens (2009, 2013) explicitly model intention-based social preferences under asymmetric information. Models within the framework of Levine (1998) rely on asymmetric information and signalling to generate reciprocity via type-dependent preferences.

[2] A large experimental literature has provided evidence for the general importance of intentions for behavior, see e.g. Blount (1995) for an early and Falk et al. (2008) for a more recent contribution. However, already Blount (1995) has pointed out that "[f]indings on attributions and the lability of preferences in social context are particularly applicable to the relationship between games of incomplete and imperfect information" (p. 142f).

[3] See e.g. Bergemann and Morris (2005) and Jehiel et al. (2006), and the discussion therein. Filiz-Ozbay and Ozbay (2007) argue that a psychological concern to avoid anticipated material ex-post regret can lead to overbidding in auctions.

type-invariance to eliminate psychological regret, as it renders intentions fully transparent in the first place. Our notion of ex-post fairness equilibrium rules out material ex-post regret at the same time. The mechanisms that we construct align individual and social objectives both on the interim and on the ex-post stage.

In spite of the strong demands implied by our notions of robustness, we show that materially Pareto efficient social choice functions can be implemented whenever they provide insurance to the agents. This requires that the expected material payoff of any player i is invariant to the private information of any other player j, where expectation is taken over the private information of all players except j.[4] In a direct revelation mechanism, the insurance property would imply that j cannot affect the payoff of i if all players except j were to tell the truth (Bierbrauer and Netzer, 2014; Bartling and Netzer, 2014). This property provides the basis for type-invariant kindness: It implies that truth-telling in the direct mechanism is associated with zero kindness for all types of all players, as nobody has the option to make anyone else better or worse off. We then exploit the reference-dependence of intention-based social preferences and augment the direct mechanism by messages which trigger additional budget-balanced transfers, but remain unused in equilibrium. This allows us to manipulate reference points and increase kindness to levels that guarantee truth-telling of all players, by turning them into maximizers of the sum of expected material payoffs, without violating the property of type-invariance. If players become informed about their opponents' types ex-post, kindness perceptions would not change and truth-telling would still be a best response, as players remain maximizers of the sum of ex-post material payoffs.

We also address the issue of voluntary participation, by considering mechanisms that give all players the right to enforce some status-quo allocation. Such veto rights complicate the construction of type-invariant kindness, as their execution might be kindness-relevant for some but not for other player types. Veto rights can thereby compromise the property of type-invariant kindness despite the insurance property of a social choice function. However, we show that such concerns can be addressed by a more complicated construction of a mechanism that uses different out-of-equilibrium transfers depending on whether or not a player profits in expectation from the execution of the veto right.

Our results complement those by Bierbrauer and Netzer (2014), who first modelled intention-based social preferences in a mechanism design framework.[5] They provide suffi-

[4] Our notion of insurance is related to similar concepts in the literature on auctions and mechanism design with risk averse bidders (Maskin and Riley, 1984) or under ambiguity (Bose et al., 2006; Bodoh-Creed, 2012). We will discuss this in Section 1.4 below.

[5] The growing literature on behavioral mechanism design has also investigated procedural motives (Glazer and Rubinstein, 1998), robustness to non-equilibrium behavior (Eliaz, 2002), honesty (e.g. Alger and Renault, 2006), state-dependent and endogenous preferences (e.g. Bowles and Polanía-Reyes, 2012; Antler, 2014), level-k reasoning (Crawford et al., 2009), learning (e.g. Mathevet, 2010; Cabrales and Serrano, 2011), lack of common knowlegde of rationality (Renou and Schlag, 2011), irrational choice functions (e.g. de Clippel, 2014) and loss aversion (Eisenhuth, 2012). Distributional preferences have been investigated by Desiraju and Sappington (2007), Kucuksenel (2012), von Siemens (2011) and Tang and Sandholm (2012). Jehiel and Moldovanu (2006) survey the large literature on externalities in mechanism design. De Marco and Immordino (2012, 2013) and Bassi et al. (2014) examine reciprocity in a team design problem, based on a model that differs from Rabin (1993) and does not exhibit reference-dependence.

cient conditions such that any materially Pareto efficient social choice function can be (voluntarily) implemented in (ex-ante or ex-interim) Bayes-Nash fairness equilibrium. Their argument rests on a manipulation of equitable payoffs very similar to the one employed here, but it does not guarantee the property of type-invariant kindness, which is central to the present chapter. The insurance property is used by Bierbrauer and Netzer (2014) to guarantee robustness of implementation results with respect to non-selfish motives of the agents.[6] The analysis in this chapter combines these arguments about robustness and about the possibility to manipulate social preferences by choice of a mechanism. Bierbrauer et al. (2014) also investigate ex-post implementation with social preferences, but, in contrast to our approach, they do not work with a specific model of social preferences. Instead they impose the insurance property as a constraint on an otherwise standard ex-post implementation problem in a bilateral trade setting.

The remainder of the chapter is organized as follows. The general framework is introduced in Section 1.2. Section 1.3 discusses our notions of robustness. The main results are presented in Section 1.4. Section 1.5 presents an application to a public goods example, and Section 1.6 concludes.

1.2 General Framework

1.2.1 Environment

The analysis builds on the formal framework of Bierbrauer and Netzer (2014). We fix an environment $E = [I, A, (\Theta_i, \pi_i)_{i \in I}, p]$, where $I = \{1, \ldots, n\}$ is a set of agents, A is a set of feasible allocations, Θ_i is a finite set of types for agent i, π_i denotes the material payoff function for agent i, and p represents a probability distribution with support $\Theta = \times_{i \in I} \Theta_i$. We employ the notation $p(\theta_i)$ and $p(\theta_{-i})$ for marginal distributions with respect to the types of subsets of agents.

As for material payoffs, we consider a quasilinear environment with independent private values. Formally, $A = Q \times T$ where Q represents a set of possible decisions and $T = \{(t_1, \ldots, t_n) \in \mathbb{R}^n | \sum_{i \in I} t_i \leq 0\}$ the set of feasible transfers. Each player's material payoff is a function $\pi_i : A \times \Theta_i \to \mathbb{R}$, given by $\pi_i(a, \theta_i) = v_i(q, \theta_i) + t_i$. Types are independent, so $p(\theta) = \prod_{i \in I} p(\theta_i)$.

A social choice function f is a mapping $f : \Theta \to A$. When referring to its specific parts, we employ the notation $f = (q^f, t_1^f, \ldots, t_n^f)$. A social choice function is materially Pareto efficient if (i) its decision rule q^f is value maximizing, $q^f(\theta) \in \arg\max_{q \in Q} \sum_{i \in I} v_i(q, \theta_i)$ for all $\theta \in \Theta$, and (ii) the transfer scheme is ex-post budget balanced, $\sum_{i \in I} t_i^f(\theta) = 0$ for all $\theta \in \Theta$. Throughout the chapter, we will restrict attention to investigating the implementability of efficient social choice functions.

[6] Bartling and Netzer (2014) investigate the insurance property in an auction setting and provide experimental evidence in favour of the theoretical robustness prediction.

1.2.2 Mechanism

A mechanism $\Phi = [M_1, \ldots, M_n, g]$ prescribes a finite set of messages M_i for every agent $i \in I$, and an outcome function $g : M \rightarrow A$ where $M = \times_{i \in I} M_i$. When referring to specific parts of the outcome function, we use the notation $g = (q^g, t_1^g, \ldots, t_n^g)$.

A mechanism Φ and the environment E jointly induce a Bayesian game, where player i's pure strategy is a function $s_i : \Theta_i \rightarrow M_i$. Denote by S_i the set of all pure strategies for player i. Let the first-order point belief of player i about player j's strategy be denoted by $s_{ij}^b \in S_j$. A complete first-order belief profile of player i is denoted by $s_i^b = (s_{ij}^b)_{j \neq i} \in S_i^b = \times_{j \neq i} S_j$. A second-order point belief of player i concerning j's first-order point belief about player k's strategy is denoted by $s_{ijk}^{bb} \in S_k$. Player i's second-order belief about j's complete first-order belief profile shall be denoted by $s_{ij}^{bb} = (s_{ijk}^{bb})_{k \neq j} \in S_j^b$. Finally, a complete second-order belief profile of player i is $s_i^{bb} = (s_{ij}^{bb})_{j \neq i} \in S_i^{bb} = \times_{j \neq i} S_j^b$.

1.2.3 Utility

We first presume every player to submit his message at the *interim* stage, where each player is informed about the own type θ_i while, at the same time, remains uninformed about the realization of the other players' types θ_{-i}. We denote the interim expected material payoff of player i from submitting m_i, given type θ_i and belief s_i^b about the other players' strategies, as

$$\Pi_i(m_i, s_i^b | \theta_i) = \mathbb{E}_{\theta_{-i}} \left[\pi_i(g(m_i, s_i^b(\theta_{-i})), \theta_i) \right].$$

Analogously, we let

$$\Pi_j(m_i, s_i^b) = \mathbb{E}_{\theta_{-i}} \left[\pi_j(g(m_i, s_i^b(\theta_{-i})), \theta_j) \right]$$

denote the material payoff that i expects to give to j when sending message m_i, given belief s_i^b.

Next, we follow the definition of interim utility proposed by Bierbrauer and Netzer (2014, Appendix B), which translates the concept of Rabin (1993) to Bayesian games. Accordingly, in addition to material payoffs $\Pi_i(m_i, s_i^b | \theta_i)$, each player is motivated by psychological reciprocity payoffs. We denote player i's interim belief about his kindness towards player j by $\kappa_{ij}(m_i, s_i^b | \theta_i)$, and his interim belief about j's kindness towards himself by $\lambda_{ji}(s_{ij}^b, s_{ij}^{bb})$. Below we will define κ_{ij} and λ_{ji} formally, in a way such that these terms take on positive values if associated with kind behavior and negative values if associated with unkind behavior. Reciprocity is captured by the assumption that mutual kindness as well as mutual unkindness increase interim utility:

$$U_i(m_i, s_i^b, s_i^{bb} | \theta_i) = \Pi_i(m_i, s_i^b | \theta_i) + \sum_{j \neq i} y_{ij} \kappa_{ij}(m_i, s_i^b | \theta_i) \lambda_{ji}(s_{ij}^b, s_{ij}^{bb}),$$

where $y_{ij} \geq 0$ indicates the degree to which other-regarding concerns matter for individual i in relation to individual j. We will also write $y_i = (y_{ij})_{j \neq i}$ and $y = (y_i)_{i \in I}$.

1.2.4 Kindness

We measure interim kindness of player i towards some other player j as the difference between the expected material payoff which player i believes to give to player j and the *equitable payoff*, $\Pi_j^{e_i}$, the reference point for the evaluation of kindness. In other words, we presume that, given his belief s_i^b, type θ_i of player i believes to be kind (unkind) towards j if his message m_i yields a higher (lower) material payoff for j than equitable. Formally,

$$\kappa_{ij}(m_i, s_i^b|\theta_i) = \Pi_j(m_i, s_i^b) - \Pi_j^{e_i}(s_i^b|\theta_i).^7$$

The equitable payoff equals a value in between the largest and smallest payoff that type θ_i of player i can give to player j, conditional on belief s_i^b:

$$\Pi_j^{e_i}(s_i^b|\theta_i) = \alpha \left[\max_{m_i \in M_i} \Pi_j(m_i, s_i^b) \right] + (1-\alpha) \left[\min_{m_i \in E_{ij}(s_i^b|\theta_i)} \Pi_j(m_i, s_i^b) \right]$$

for some $\alpha \in (0,1)$. The set of messages relevant for the minimization contains only messages which induce bilaterally Pareto efficient payoff pairs: $E_{ij}(s_i^b|\theta_i) = \{m_i \in M_i | \nexists\, m_i' \in M_i$ with $\Pi_i(m_i', s_i^b|\theta_i) \geq \Pi_i(m_i, s_i^b|\theta_i)$ and $\Pi_j(m_i', s_i^b) \geq \Pi_j(m_i, s_i^b)$, with at least one strict inequality$\}$. This assumption guarantees that messages which hurt player j without benefitting player i do not influence the equitable payoff and hence the kindness perception of i's behavior.

The message m_i submitted by type θ_i of player i determines his intended interim kindness $\kappa_{ij}(m_i, s_i^b|\theta_i)$ towards player j, given his belief s_i^b. Suppose player i knew j's type at the interim stage. Given beliefs s_{ij}^b and s_{ij}^{bb}, he could then derive a belief $\kappa_{ji}(s_{ij}^b(\theta_j), s_{ij}^{bb}|\theta_j)$ about j's intended kindness towards himself. On the interim stage, however, player i is uninformed about θ_j and therefore cannot put himself into player j's interim shoes in order to figure out the intended kindness precisely. Following Bierbrauer and Netzer (2014), we first proceed under the assumption that player i forms his belief about j's kindness by taking the expectation over θ_j,

$$\lambda_{ji}(s_{ij}^b, s_{ij}^{bb}) = \sum_{\theta_j \in \Theta_j} p(\theta_j) \kappa_{ji}(s_{ij}^b(\theta_j), s_{ij}^{bb}|\theta_j). \tag{1.1}$$

1.2.5 Equilibrium

We can now provide a definition of interim fairness equilibrium, adapted to the present notation from Bierbrauer and Netzer (2014, p. 40).

[7] Bierbrauer and Netzer (2014, Appendix B) do not explicitly specify interim kindness intentions. Instead, they provide a condition on interim kindness such that their concept of Bayes-Nash fairness equilibrium, which is based on an ex-ante perspective, is identical to their notion of interim fairness equilibrium, which takes the interim perspective. In contrast to our definition of interim kindness, Bierbrauer and Netzer (2014) allow for an upper and lower bound on ex-ante kindness.

Definition 1.1. *An interim fairness equilibrium (IFE) is a profile s^* such that, for all $i \in I$,*
(i) $s_i^*(\theta_i) \in \arg\max_{m_i \in M_i} U_i(m_i, s_i^b, s_i^{bb}|\theta_i)$ *for all* $\theta_i \in \Theta_i$*, and*
(ii) $s_{ij}^b = s_j^*$ *for all* $j \neq i$*, and*
(iii) $s_{ijk}^{bb} = s_k^*$ *for all* $j \neq i, k \neq j$*.*

A social choice function f is implementable in IFE if there exists a mechanism Φ with an IFE s^* such that $g(s^*(\theta)) = f(\theta)$ for all $\theta \in \Theta$.

1.3 Notions of Robustness

1.3.1 Type-Invariance of Kindness

Picking up on the discussion in the introduction, the fact that kindness $\kappa_{ij}(s_i^*(\theta_i), s_{-i}^*|\theta_i)$ in IFE will generally depend on θ_i may be troublesome. A mechanism designer might not be confident that (1.1) correctly reflects the way players form beliefs about the others' intentions on the interim stage. We therefore refine the concept of IFE in order to guarantee that interim equilibrium intentions are fully transparent to every player, despite the presence of private information. In particular, we require a strategy profile s^* not only to be an IFE but also to generate type-invariant equilibrium kindness values. Formally,

$$k_{ij}(s_i^*(\theta_i), s_{-i}^*|\theta_i) = k_{ij}(s_i^*(\tilde{\theta}_i), s_{-i}^*|\tilde{\theta}_i) =: k_{ij}(s^*) \text{ for all } \theta_i, \tilde{\theta}_i \in \Theta_i \text{ and } i, j \in I, j \neq i. \tag{1.2}$$

An IFE s^* that additionally satisfies condition (1.2) is called *interim fairness equilibrium with type-invariant kindness* (IFE-TI). Implementability in IFE-TI is defined accordingly.

For instance, suppose beliefs about intentions and kindness are formed in a way that deviates from the simple expectation formulation in (1.1). Let

$$\Delta_{ji}(s_{ij}^b, s_{ij}^{bb}) = \left[\min_{\theta_j \in \Theta_j} \kappa_{ji}(s_{ij}^b(\theta_j), s_{ij}^{bb}|\theta_j),\ \max_{\theta_j \in \Theta_j} \kappa_{ji}(s_{ij}^b(\theta_j), s_{ij}^{bb}|\theta_j)\right]$$

denote the interval spanned by the smallest and the largest values of j's equilibrium interim kindness towards i. An IFE-TI remains an equilibrium if we replace (1.1) by the assumption that $\lambda_{ji}(s_{ij}^b, s_{ij}^{bb}) \in \Delta_{ji}(s_{ij}^b, s_{ij}^{bb})$, without specifying any additional details. For instance, if players use arbitrary weights $w_{ij}(\theta_j) \geq 0$ with $\sum_{\theta_j \in \Theta_j} w_{ij}(\theta_j) = 1$ instead of $p(\theta_j)$ to calculate λ_{ji}, or even focus on one of the extremes such as the least kind type of the opponent, IFE-TI remains robust as it collapses Δ_{ji} to a single value.

1.3.2 Ex-Post Fairness Implementation

Our notion of ex-post fairness implementation shall provide robustness in the sense that every player should stick to his interim decision even if he were informed ex-post about the others' types. Such additional information would allow each player i to update his beliefs about each opponent j's actual interim intention, thus moving from $\lambda_{ji}(s_{ij}^b, s_{ij}^{bb})$ to

$\kappa_{ji}(s^b_{ij}(\theta_j), s^{bb}_{ij}|\theta_j)$. Notice that κ_{ji} is still based on an expectation over θ_{-j}, which reflects the information under which type θ_j actually made his choice. We then define player i's ex-post utility as

$$U_i(m_i, s^b_i, s^{bb}_i|\theta) = \Pi_i(m_i, s^b_i|\theta) + \sum_{j\neq i} y_{ij}\kappa_{ij}(m_i, s^b_i|\theta)\kappa_{ji}(s^b_{ij}(\theta_j), s^{bb}_{ij}|\theta_j), \tag{1.3}$$

where $\Pi_i(m_i, s^b_i|\theta) = \pi_i(g(m_i, s^b_i(\theta_{-i})), \theta_i)$ are the ex-post material payoffs and $\kappa_{ij}(m_i, s^b_i|\theta)$ is i's ex-post kindness toward j, defined as the difference between $\Pi_j(m_i, s^b_i|\theta)$ $= \pi_j(g(m_i, s^b_i(\theta_{-i})), \theta_j)$ and some equitable payoff.[8]

We say that a social choice function f is ex-post fairness implementable if it is implementable in an IFE-TI s^* and if, for each player i and every type profile $\theta \in \Theta$, $s^*_i(\theta_i)$ still constitutes a best-response in terms of ex-post utility (1.3), given beliefs fixed on s^*. This definition captures the above stated robustness concern, since every player would stick to his equilibrium interim decision even if he observed the others' private information on the ex-post stage. Now observe that implementation of a materially Pareto efficient social choice function in an IFE-TI s^* implies ex-post fairness implementation if s^* gives rise to the kindness values $\kappa_{ji}(s^*_j(\theta_j), s^*_{-j}|\theta_j) = 1/y_{ij}$ for all pairs of players.[9] To see the point in more detail, substitute these kindness values into (1.3) and note that $s^*_i(\theta_i)$ then constitutes a best response in ex-post utility terms if and only if it maximizes the sum of all players' ex-post material payoffs. By presupposition, strategy profile s^* results in an efficient, i.e., payoff-sum maximizing allocation $g(s^*(\theta)) = f(\theta)$ for any type profile $\theta \in \Theta$, so that this is indeed the case. We will address in the following section whether and how these conditions can be achieved.

The proposed notion of ex-post fairness implementation is still restrictive. In particular, the kindness that i attributes to θ_j's interim behavior in (1.3) corresponds to the true kindness of θ_j in a mechanism where all choices are made on the interim stage. If a mechanism systematically grants players the right to reconsider their decisions ex-post, and this is anticipated on the interim stage, then the interim kindness of a given message might be different in the first place. An analysis of games with multiple stages of announcements is currently impeded by the lack of a theory of intentions for general extensive-form games. Our main intuition for ex-post fairness implementation parallels the intuition for ex-post Nash equilibrium provided by Crémer and McLean (1985, p. 349): "Of course, in our model, a bidder can never observe the types of the other bidders. Thus, the concept of ex post Nash equilibrium corresponds to the following reasoning by agent i. "If I believe that the other bidders are using [their equilibrium strategies], then even if I observed their actions, I would have no incentive to change mine"".[10] Besides this general intuition, ex-

[8] Since the equitable payoff does not play a role for the present purposes, we omit its exact specification.

[9] Any strategy profile that results in a materially Pareto efficient social choice function and satisfies this particular condition on type-invariant kindness values must in fact be an IFE-TI, as all players are then maximizing the sum of expected material payoffs at the interim stage. However, not every IFE-TI exhibits these particular values. See the discussion following Proposition 1 below, and the example in Section 1.5.

[10] Crémer and McLean's concept corresponds to the earlier notion of uniform equilibrium proposed by

post fairness implementation can also be appropriate when the anticipation of regret affects interim decisions, in the spirit of Filiz-Ozbay and Ozbay (2007). Finally, our construction also applies when agents do in fact observe the others' types and can revise their decisions ex-post, but do not anticipate this on the interim stage.

1.4 General Results

1.4.1 Insurance and Implementation

A concept which will be very important is the *insurance property.* Intuitively, a social choice function gives rise to this property if each player i is insured against the realization of the type (or the report in a direct mechanism) of any other player j, provided an expectation is taken over the types of all other players (or provided that all other players report truthfully). Hence unilateral deviations from truth-telling in the direct mechanism will not affect any other player's payoff when the insurance property is satisfied (Bierbrauer and Netzer, 2014; Bartling and Netzer, 2014).

Definition 1.2. *Given an environment E and social choice function f, the insurance property holds if for each $i \in I$ there exists $P_i \in \mathbb{R}$ such that*

$$\mathbb{E}_{\theta_{-j}}\left[\pi_i\big(f(\tilde{\theta}_j, \theta_{-j}), \theta_i\big)\right] = P_i$$

for all $j \neq i$ and $\tilde{\theta}_j \in \Theta_j$.

Related notions of insurance exist in the literature on optimal auctions with risk averse bidders (e.g. Maskin and Riley, 1984) or with ambiguity (e.g. Bose et al., 2006). Maskin and Riley (1984) define a *perfect insurance* auction (p. 1491) where each bidder's payment is deterministic and depends only on the own type and the event of winning or losing the auction, with marginal utilities of income being equated across these two cases. In our framework with material payoffs that are linear in transfers, this is satisfied by a large class of mechanisms, such as first-price or all-pay auctions which do not satisfy our insurance property based on overall payoffs.[11] Bose et al. (2006) define a *full insurance* mechanism (p. 416) where each bidder's ex-post payoff depends only on the own type, and they show that full insurance is optimal with ambiguity averse bidders.[12] The property of full insurance is stronger than our insurance property, as we require invariance of payoffs with respect to another player's type only from an ex-ante expected perspective. We can now state our first main result.

d'Aspremont and Gerard-Varet (1979), which builds on the concept of "complete ignorance" (see e.g. Luce and Raiffa, 1957).

[11] See Eisenhuth (2012) and Eisenhuth and Ewers (2012) for an analysis of such mechanisms with loss averse bidders. Maskin and Riley (1984) show that a perfect insurance auction will typically not be optimal with risk averse bidders, when the auctioneer can use risk to relax incentive constraints.

[12] Perfect and full insurance coincide for certain classes of risk preferences, see Bose et al. (2006) for a discussion.

Proposition 1.1. *Assume that $y \in]0, \infty[^{n(n-1)}$. If a social choice function f is materially Pareto efficient and the insurance property is satisfied, then f is implementable in an IFE-TI s^* in which $\kappa_{ij}(s^*) = 1/y_{ji}$ holds for all pairs of players.*

Proof of Proposition 1.1. We first prove the result for $n = 2$. We comment on the case where $n > 2$ afterwards. Throughout, we fix a social choice function f that is efficient and we suppose that the insurance property holds. We also assume $y_{12}, y_{21} > 0$, and we proceed in two steps. First, we construct a specific mechanism Φ for f. Second, we show that Φ has an IFE-TI in which f is realized with the desired kindness levels.

Step 1. Construct mechanism $\Phi = [M_1, M_2, g]$ as follows. For both $i = 1, 2$ we let $M_i = \Theta_i \times \{0, 1\}$, so that a message $m_i = (\eta_i, \gamma_i) \in M_i$ contains an announced type $\eta_i \in \Theta_i$ and an announced number $\gamma_i \in \{0, 1\}$. Given a message profile $m = (m_1, m_2) \in M$, we also write $\eta = (\eta_1, \eta_2) \in \Theta$ for the profile of announced types, and $\gamma = (\gamma_1, \gamma_2) \in \{0, 1\}^2$ for the profile of announced numbers. The outcome function g is defined as follows. For all $m \in M$, let $q^g(m) = q^f(\eta)$, i.e. only the announced types η matter for the decision rule, which follows f. For all $m \in M$, $i = 1, 2$ and $j \neq i$, let $t_i^g(m) = t_i^f(\eta) + r_i(\gamma)$, where

$$r_i(\gamma) = \begin{cases} +e_{ij} & \text{if } \gamma_i = 1, \gamma_j = 0, \\ -e_{ji} & \text{if } \gamma_i = 0, \gamma_j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, transfers also depend on the announced types η according to f, plus an additional term that depends on the announced numbers γ. If i announces $\gamma_i = 1$ and j announces $\gamma_j = 0$, then i takes an additional amount of e_{ij} from j, and vice versa. In the following, use of i and j always presumes $j \neq i$. Since f is efficient and the additional transfers r_i always sum to zero across players, mechanism Φ is budget balanced for all profiles $m \in M$. As long as the announcements satisfy $\gamma_i = \gamma_j = 0$, the mechanism corresponds to a direct mechanism for f.

Step 2. Consider strategy profile $s^T = (s_1^T, s_2^T)$ where $s_i^T(\theta_i) = (\theta_i, 0)$ for all $\theta_i \in \Theta_i$. The profile s^T corresponds to truth-telling in a direct mechanism. Under s^T we have $g(s^T(\theta)) = f(\theta)$ for all $\theta \in \Theta$, so that f is realized. We will show that, for appropriately chosen values of e_{12} and e_{21}, strategy profile s^T is an IFE-TI of Φ, and the desired kindness levels arise. In the hypothetical equilibrium s^T, beliefs are correct: $s_i^b = s_j^{bb} = s_j^T$ for $i = 1, 2$.

We first derive the kindness term $\kappa_{ij}((\theta_i, 0), s_j^T | \theta_i)$. Given the definition of Φ, choice of m_i by player i induces the following pair of payoffs:

$$\left[\Pi_i(m_i, s_j^T | \theta_i), \Pi_j(m_i, s_j^T)\right] = \begin{cases} \left[P_i(\eta_i, \theta_i), P_j\right] & \text{for } m_i = (\eta_i, 0) \text{ with } \eta_i \in \Theta_i, \\ \left[P_i(\eta_i, \theta_i) + e_{ij}, P_j - e_{ij}\right] & \text{for } m_i = (\eta_i, 1) \text{ with } \eta_i \in \Theta_i, \end{cases}$$

where $P_i(\eta_i, \theta_i) = \mathbb{E}_{\theta_j}[\pi_i(f(\eta_i, \theta_j), \theta_i)]$, and $P_j = \mathbb{E}_{\theta_j}[\pi_j(f(\eta_j, \theta_j), \theta_j)]$ is a constant because of the insurance property.

Fix $e_{ij} = 1/[(1-\alpha)y_{ji}]$, implying $\max_{m_i \in M_i} \Pi_j(m_i, s_j^T) = P_j$. Let $\eta_i^* \in \arg\max_{\eta_i \in \Theta_i} P_i(\eta_i, \theta_i)$. Since any message $(\eta_i^*, 1)$ simultaneously maximizes $\Pi_i(m_i, s_j^T|\theta_i)$ and minimizes $\Pi_j(m_i, s_j^T)$, we have $(\eta_i^*, 1) \in E_{ij}(s_j^T|\theta_i)$, and thus $\min_{m_i \in E_{ij}(s_j^T|\theta_i)} \Pi_j(m_i, s_j^T) = P_j - 1/[(1-\alpha)y_{ji}]$. It follows that $\Pi_j^{e_i}(s_j^T|\theta_i) = P_j - 1/y_{ji}$ and therefore $\kappa_{ij}((\theta_i, 0), s_j^T|\theta_i) = 1/y_{ji}$ as required. Replicating the same argument for player j and $e_{ji} = 1/[(1-\alpha)y_{ij}]$ yields the type-invariant kindness $\kappa_{ji}((\theta_j, 0), s_i^T|\theta_j) = 1/y_{ij}$. We thus have $\lambda_{ij}(s^T) = 1/y_{ji}$ and $\lambda_{ji}(s^T) = 1/y_{ij}$ in the hypothetical equilibrium.

Consider now player i's interim utility for any $\theta_i \in \Theta_i$:

$$U_i(m_i, s_j^T, s_i^T|\theta_i) = \Pi_i(m_i, s_j^T|\theta_i) + y_{ij}\kappa_{ij}(m_i, s_j^T|\theta_i) \cdot (1/y_{ij}).$$

Omitting terms that do not depend on m_i, $m_i = (\theta_i, 0)$ is a maximizer of this expression if and only if it is a maximizer of

$$\Pi_i(m_i, s_j^T|\theta_i) + \Pi_j(m_i, s_j^T) = \mathbb{E}_{\theta_j}\left[v_i(q^g(m_i, s_j^T(\theta_j)), \theta_i) + v_j(q^g(m_i, s_j^T(\theta_j)), \theta_j)\right],$$

where the equality holds due to budget balance of Φ. Since f is efficient, so that q^f is value maximizing, $m_i = (\theta_i, 0)$ is a solution to the interim utility maximization problem, for any $\theta_i \in \Theta_i$. Replicating the argument for player j, we can conclude that s^T is an IFE-TI.

The case of $n > 2$. The arguments for $n = 2$ can be generalized to the case of $n > 2$, by defining message sets $M_i = \Theta_i \times [\{0\} \cup (I \backslash \{i\})]$. The announcement of $m_i = (\eta_i, \gamma_i)$ corresponds to the announcement of type η_i in a direct mechanism, but player i obtains an additional transfer e_{ij} from player j if and only if $\gamma_i = j$ and $\gamma_k = 0$ for all $k \neq i$. The above arguments can then be applied analogously for each pair of players, and bilateral type-invariant kindness of truth-telling $s_i^T(\theta_i) = (\theta_i, 0)$ can be adjusted by choice of the additional transfers payments so that each player's goal becomes the maximization of the sum of material payoffs. □

The mechanism that we construct in the proof of Proposition 1.1 works like a direct mechanism, where each player announces a type, but with the new feature that each player i can claim an additional payment of e_{ij} from any opponent j.[13] Truthful revelation s^* without claiming such payments then becomes kind behavior. We show that the (budget-balanced) payments e_{ij} can be adjusted so that the type-invariant interim kindness $\kappa_{ij}(s^*) = 1/y_{ji}$ is achieved for each pair of players. The *insurance property* is crucial for this to be possible. It implies that the realized and revealed type is irrelevant for kindness; all that matters is the fact that no additional payment is claimed. Each player then becomes a maximizer of the sum of expected material payoffs on the interim stage. Since the social choice function f is *efficient*, truth-telling is then a best-response to truth-telling of the opponents,

[13] Bierbrauer and Netzer (2014) refer to such mechanism as a "mechanism with a button". It is isomorphic to an *augmented revelation mechanism* (Mookherjee and Reichelstein 1994). The reason why it is formally not an augmented revelation mechanism is that we define message sets with a product structure, $M_i = \Theta_i \times D_i$ for some set D_i, instead of defining it as a union $M_i = \Theta_i \cup D_i$ so that $\Theta_i \subseteq M_i$. Saran (2011) discusses the validity of the revelation principle for general menu-dependent preferences.

which implies that the mechanism implements f in IFE-TI. Due to the specific values of equilibrium kindness, our arguments from Section 1.3.2 imply that it also implements f in ex-post fairness equilibrium.

If only implementation in IFE-TI but not ex-post fairness implementation was required, a much simpler construction would suffice. In fact, the direct revelation mechanism implements any efficient social choice function with the insurance property in IFE-TI. This is true because efficiency and insurance jointly imply standard Bayesian incentive-compatibility (see Lemma 1.1 below) and Bayesian incentive-compatibility and the insurance property jointly imply that f is implemented by a fairness equilibrium with (type-invariant) kindness levels of zero in the direct mechanism (see Bierbrauer and Netzer, 2014).

Lemma 1.1. *If f is materially Pareto efficient and satisfies the insurance property, then f is Bayesian incentive-compatible.*

Proof of Lemma 1.1. From material Pareto efficiency of f it follows that

$$\sum_{i=1}^{n} \pi_i(f(\theta_j, \theta_{-j}), \theta_i) \geq \sum_{i=1}^{n} \pi_i(f(\hat{\theta}_j, \theta_{-j}), \theta_i)$$

for all $j \in I$, $(\theta_j, \theta_{-j}) \in \Theta$ and $\hat{\theta}_j \in \Theta_j$. Taking expectation with respect to θ_{-j}, this becomes

$$\sum_{i=1}^{n} \mathbb{E}_{\theta_{-j}}[\pi_i(f(\theta_j, \theta_{-j}), \theta_i)] \geq \sum_{i=1}^{n} \mathbb{E}_{\theta_{-j}}[\pi_i(f(\hat{\theta}_j, \theta_{-j}), \theta_i)].$$

Due to the insurance property of f, we have that

$$\mathbb{E}_{\theta_{-j}}[\pi_i(f(\tilde{\theta}_j, \theta_{-j}), \theta_i)] = P_i$$

is independent of $\tilde{\theta}_j$ for all agents $i \neq j$, so that we can simplify the inequality to

$$\mathbb{E}_{\theta_{-j}}[\pi_j(f(\theta_j, \theta_{-j}), \theta_j)] \geq \mathbb{E}_{\theta_{-j}}[\pi_j(f(\hat{\theta}_j, \theta_{-j}), \theta_j)],$$

which is the conventional Bayesian incentive-compatibility condition. □

As the next section shows, even IFE-TI implementation (without the additional requirement of ex-post implementability) will become more difficult when voluntary participation is required.

1.4.2 Voluntary Participation

The analysis in the previous section ignored the question whether or not some type of some player would prefer to opt out of the mechanism (see Myerson and Satterthwaite, 1983, for the classical impossibility result). To show that *voluntary participation* can be guaranteed as well, we now require that the mechanism used to implement f admits veto rights: every player must have a message which enforces a fixed status quo allocation

$\bar{a} = (\bar{q}, \bar{t}_1, \ldots, \bar{t}_n) \in A$. We assume that $\bar{a}$ is budget balanced, $\sum_{i \in I} \bar{t}_i = 0$, but allow it to be chosen arbitrarily otherwise.[14] If IFE-TI implementation of a social choice function f is possible in such a mechanism, we say that f is voluntarily implementable in IFE-TI.

Voluntary implementation raises several novel issues compared to the previous section. First, the direct mechanism with veto rights no longer implements f in IFE with zero kindness, despite efficiency and the insurance property, because some types of some players might prefer to opt out of the mechanism. Second, veto rights generally complicate the problem of achieving type-invariant kindness. Execution of the veto may induce bilaterally Pareto efficient payoff pairs for some but not for other types, so that the veto is relevant for the computation of equitable payoffs in the former but not in the latter case. Equitable payoffs and kindness can therefore vary with the realized type despite the insurance property of f, because the insurance property constrains payoffs derived from type reports but not from the exercise of a veto. Finally, truth-telling in a direct mechanism with veto rights goes along with different kindness values than in a direct mechanism without veto rights. Hence our construction of off-equilibrium payments must be different, and, in particular, it can become necessary to decrease equilibrium kindness. Nevertheless, we can establish the following result:

Proposition 1.2. *Assume that $y \in]0, \infty[^{n(n-1)}$. If a social choice function f is materially Pareto efficient and the insurance property is satisfied, then f is voluntarily implementable in an IFE-TI s^* in which $\kappa_{ij}(s^*) = 1/y_{ji}$ holds for all pairs of players.*

Proof of Proposition 1.2. As before, we first prove the result for $n = 2$ and comment on the case $n > 2$ afterwards. We fix a social choice function f that is efficient and we suppose that the insurance property holds. We also fix an arbitrary budget balanced status quo $\bar{a} = (\bar{q}, \bar{t}_1, \bar{t}_2) \in A$. We proceed in two steps. First, we construct a mechanism Φ for f which admits veto rights. Second, we show that Φ has an IFE-TI in which f is realized with the desired kindness levels.

Step 1. Construct mechanism $\Phi = [M_1, M_2, g]$ as follows. Let $M_i = (\Theta_i \cup \{\nu\}) \times \{0, 1\}$, so that a message $m_i = (\eta_i, \gamma_i) \in M_i$ again comprises two components. First, $\eta_i \in \Theta_i \cup \{\nu\}$ allows player i to report either a type from Θ_i or to exercise a veto ν. Second, player i announces a number $\gamma_i \in \{0, 1\}$. Given a profile $m = (m_1, m_2) \in M$, we again write $\eta = (\eta_1, \eta_2)$ and $\gamma = (\gamma_1, \gamma_2)$. The outcome function g is defined differently for two cases. First, if m has $\eta_i = \nu$ for at least one $i = 1, 2$, we let $q^g(m) = \bar{q}$ and $t_i^g(m) = \bar{t}_i + \bar{r}_i(\gamma)$, where

$$\bar{r}_i(\gamma) = \begin{cases} +d_{ij} & \text{if } \gamma_i = 1, \gamma_j = 0, \\ -d_{ji} & \text{if } \gamma_i = 0, \gamma_j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Hence allocation $\bar{a}$ is chosen, with possible additional transfers depending on γ. Second, if

[14] Our assumption implies that the mechanism remains budget balanced out-of-equilibrium, which simplifies the proof. Our result would continue to hold, however, if $\sum_{i \in I} \bar{t}_i < 0$ was true.

m has $\eta \in \Theta$, we let $q^g(m) = q^f(\eta)$ and $t_i^g(m) = t_i^f(\eta) + r_i(\gamma)$, where

$$r_i(\gamma) = \begin{cases} +e_{ij} & \text{if } \gamma_i = 1, \gamma_j = 0, \\ -e_{ji} & \text{if } \gamma_i = 0, \gamma_j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the outcome function selects allocation $f(\eta)$ where additional transfers may occur in accordance with γ. Since f is efficient, $\bar{a}$ is budget balanced, and the additional transfers $\bar{r}_i$ and r_i always sum to zero across players, Φ is budget balanced for all profiles $m \in M$. If the announcements satisfy $\gamma_i = \gamma_j = 0$, then the allocations induced by Φ are equivalent to the allocations given by a direct mechanism with additional veto rights for every player.

Step 2. Consider $s^T = (s_1^T, s_2^T)$ where $s_i^T(\theta_i) = (\theta_i, 0)$ for all $\theta_i \in \Theta_i$, so that the veto rights remain unused and $g(s^T(\theta)) = f(\theta)$ for all $\theta \in \Theta$. We will show that s^T is an IFE-TI in which the desired kindness levels arise for appropriate values of e_{12}, e_{21}, d_{12} and d_{21}.

We first derive $\kappa_{ij}((\theta_i, 0), s_j^T|\theta_i)$. Player i with type θ_i can induce the following payoff pairs:

$$\left[\Pi_i(m_i, s_j^T|\theta_i), \Pi_j(m_i, s_j^T)\right] = \begin{cases} \left[P_i(\eta_i, \theta_i), P_j\right] & \text{for } m_i = (\eta_i, 0) \text{ with } \eta_i \in \Theta_i, \\ \left[P_i(\eta_i, \theta_i) + e_{ij}, P_j - e_{ij}\right] & \text{for } m_i = (\eta_i, 1) \text{ with } \eta_i \in \Theta_i, \\ \left[\bar{P}_i(\theta_i), \bar{P}_j\right] & \text{for } m_i = (\nu, 0), \\ \left[\bar{P}_i(\theta_i) + d_{ij}, \bar{P}_j - d_{ij}\right] & \text{for } m_i = (\nu, 1), \end{cases}$$

where $P_i(\eta_i, \theta_i)$ and P_j are defined as in the proof of Proposition 1.1, $\bar{P}_i(\theta_i) = \pi_i(\bar{a}, \theta_i)$ is player i's material payoff in $\bar{a}$ and $\bar{P}_j = \mathbb{E}_{\theta_j}[\pi_j(\bar{a}, \theta_j)]$ is player j's expected material payoff from $\bar{a}$. Define $\delta_j = \bar{P}_j - P_j$, which does not depend on any type and thus is known to the mechanism designer, who can distinguish between the following three cases:

(a): $\delta_j \geq 0$. Let $e_{ij} = [1 + \alpha y_{ji} \delta_j]/[(1-\alpha) y_{ji}]$ and $d_{ij} = [1 + y_{ji}\delta_j]/[(1-\alpha)y_{ji}]$, so $e_{ij}, d_{ij} > 0$. We obtain

$$\bar{P}_j \geq P_j > P_j - e_{ij} = \bar{P}_j - d_{ij}$$

and hence $\max_{m_i \in M_i} \Pi_j(m_i, s_j^T) = \bar{P}_j$. Player i's own payoff can be maximized only by either a message $m_i = (\eta_i^*, 1)$ for $\eta_i^* \in \arg\max_{\eta_i \in \Theta_i} P_i(\eta_i, \theta_i)$, or by message $m_i = (\nu, 1)$. Hence, one of these messages must belong to $E_{ij}(s_j^T|\theta_i)$. All of them yield the same minimal payoff for j, so $\min_{m_i \in E_{ij}(s_j^T|\theta_i)} \Pi_j(m_i, s_j^T) = \bar{P}_j - d_{ij}$. As a result, $\Pi_i^{e_i}(s_j^T|\theta_i) = P_j - 1/y_{ji}$ and $\kappa_{ij}((\theta_i, 0), s_j^T|\theta_i) = 1/y_{ji}$.

(b): $-1/[(1-\alpha)y_{ji}] < \delta_j < 0$. Let $e_{ij} = 1/[(1-\alpha)y_{ji}]$ and $d_{ij} = [1 + (1-\alpha)y_{ji}\delta_j]/[(1-\alpha)y_{ji}]$, so again $e_{ij}, d_{ij} > 0$. We obtain

$$P_j > \bar{P}_j > P_j - e_{ij} = \bar{P}_j - d_{ij}.$$

Thus, $\max_{m_i \in M_i} \Pi_j(m_i, s_j^T) = P_j$ and $\min_{m_i \in E_{ij}(s_j^T|\theta_i)} \Pi_j(m_i, s_j^T) = P_j - e_{ij}$, by the same argument as in case (a). This again implies $\Pi_i^{e_i}(s_j^T|\theta_i) = P_j - 1/y_{ji}$ and $\kappa_{ij}((\theta_i, 0), s_j^T|\theta_i) = 1/y_{ji}$.

(c): $\delta_j \leq -1/[(1-\alpha)y_{ji}]$. Let $e_{ij} = -\delta_j$ and $d_{ij} = [1 + y_{ji}\delta_j]/[\alpha y_{ji}]$, so $e_{ij} > 0$ and $d_{ij} < 0$. We obtain

$$\bar{P}_j - d_{ij} \geq P_j > \bar{P}_j = P_j - e_{ij}$$

and $\max_{m_i \in M_i} \Pi_j(m_i, s_j^T) = \bar{P}_j - d_{ij}$. Player i's own payoff can be maximized only by a message $(\eta_i^*, 1)$ where $\eta_i^* \in \arg\max_{\eta_i \in \Theta_i} P_i(\eta_i, \theta_i)$, or by message $(\nu, 0)$. Hence one of these messages must be contained in $E_{ij}(s_j^T|\theta_i)$. All of them yield the same minimal payoff for j, so that $\min_{m_i \in E_{ij}(s_j^T|\theta_i)} \Pi_j(m_i, s_j^T) = \bar{P}_j$, $\Pi_i^{e_i}(s_j^T|\theta_i) = P_j - 1/y_{ji}$, and $\kappa_{ij}((\theta_i, 0), s_j^T|\theta_i) = 1/y_{ji}$.

We have shown that $\kappa_{ij}((\theta_i, 0), s_j^T|\theta_i) = 1/y_{ji}$ can be achieved by an appropriate choice of e_{ij} and d_{ij} in any case. The remainder of the proof is analogous to the proof of Proposition 1.1.

The case of $n > 2$. As for Proposition 1.1, the arguments for $n = 2$ can be generalized to the case of $n > 2$, now using message sets $M_i = (\Theta_i \cup \{\nu\}) \times [\{0\} \cup (I \backslash \{i\})]$. □

The mechanism constructed in the proof can be interpreted as follows: We first fix the direct revelation mechanism for f and extend it with veto rights. Analogous to the construction for Proposition 1.1, we then allow each player to claim extra payments from any other player, over and above the transfers of f or the status quo allocation. The goal of these extra payments is again to manipulate kindness of truth-telling to values that turn players into maximizers of the sum of expected material payoffs. The amount of these payments depends on whether the claims are coupled with either a type report or the execution of the veto right. When enforcement of the status quo benefits or does not hurt player j too strongly (cases (a) and (b) in the proof), then the transfers that i can claim from j in addition to executing the veto are designed to give the same minimal payoff to player j as when i reports a type and claims the associated additional transfers. This minimum is therefore independent of whether or not the status quo is bilaterally Pareto efficient from the perspective of type θ_i of player i, which yields the desired type-invariance. If enforcement of the status quo hurts player j strongly (case (c) in the proof), then execution of the veto defines the minimal payoff that i can give to j. The transfers that i can claim in addition to reporting a type are tailored to induce that same minimum, which again yields type-invariance. The transfers that i can claim in addition to the veto become negative and are used to adjust the maximal payoff that i can give to j until the desired equitable payoff is obtained. This construction stabilizes equitable payoffs and therefore kindness such that it does not vary with the realized type in equilibrium.

1.5 Public Goods Example

1.5.1 Environment

In this section, we will illustrate the relevance of the insurance property and provide examples of the mechanisms used in our proofs. We work with a simple public goods application.[15] Consider an environment with two players, $I = \{1, 2\}$, and the problem of whether or not to provide a public good, $Q = \{0, 1\}$. Each player can be of either high or low type, $\Theta_i = \{\theta_i^L, \theta_i^H\}$, both of which are equally likely. Types capture the players' willingness to pay for the public good, so we have $v_i(1, \theta_i) = \theta_i - c$ and $v_i(0, \theta_i) = 0$, where $c > 0$ is the per capita cost of providing the public good, assumed to be shared equally by default. We assume

$$c < \theta_1^L < \theta_1^H \quad \text{and} \quad \theta_2^L < c < \theta_2^H,$$

which implies that player 1 would always like to have the public good provided, but player 2 only if he has the high type. We also assume that

$$\theta_1^L + \theta_2^L < 2c < \theta_1^H + \theta_2^L,$$

which implies that Pareto efficiency requires to provide the good except if $\theta = (\theta_1^L, \theta_2^L)$.

1.5.2 Expected Externality Mechanism

The expected externality mechanism (AGV) for an efficient decision rule (d'Aspremont and Gerard-Varet, 1979; Arrow, 1979) is ex-post budget balanced, hence materially Pareto efficient, and Bayesian incentive-compatible. As Bierbrauer and Netzer (2014) have shown, it always satisfies the insurance property for the case of two players. Table 1 summarizes this mechanism for our example, by stating the decision rule q^f and player 1's transfer t_1^f. The transfer for player 2 is given by $t_2^f = -t_1^f$.

Table 1.1: AGV (q^f, t_1^f).

	θ_2^L	θ_2^H
θ_1^L	$(0, (\theta_2^H - \theta_1^H)/2)$	$(1, (\theta_2^H - \theta_1^H - \theta_1^L + c)/2)$
θ_1^H	$(1, (\theta_2^H - \theta_1^H + \theta_2^L - c)/2)$	$(1, (\theta_2^H - \theta_1^H + \theta_2^L - \theta_1^L)/2)$

Ex-ante expected payoffs of the two players in the truth-telling Bayes-Nash equilibrium are

$$P_1 = \frac{1}{2}\theta_2^H + \frac{1}{4}\theta_2^L - \frac{3}{4}c \quad \text{and} \quad P_2 = \frac{1}{2}\theta_1^H + \frac{1}{4}\theta_1^L - \frac{3}{4}c.$$

Due to the insurance property, each player $i = 1, 2$ obtains the same expected payoff P_i even if the other player $j \neq i$ deviates from truth-telling to any of the other possible

[15] This example application was also used in Bierbrauer and Netzer (2011), an earlier version of Bierbrauer and Netzer (2014).

strategies. This implies that the truth-telling strategy profile is associated with type-invariant kindness levels of zero in this mechanism. Psychological concerns are therefore irrelevant to both players, and Bayesian incentive-compatibility implies that truth-telling is also an IFE-TI.

1.5.3 Ex-Post Fairness Implementation

The previous result can still be seen as a corollary of the general robustness arguments in Bierbrauer and Netzer (2014). The AGV does, however, not guarantee ex-post fairness implementation. To see why, assume that both players have followed a truth-telling strategy ex-interim, have correct beliefs about this fact, and type profile $\theta = (\theta_1^L, \theta_2^L)$ has realized. Since updating the interim kindness values to the new information still results in mutual kindness of zero, the maximization of ex-post utility (1.3) boils down to a maximization of own material ex-post payoffs for both players. The condition for player 1 wanting to deviate ex-post to the non-truthful type announcement θ_1^H becomes $(\theta_1^L - c) + (\theta_2^L - c)/2 > 0$. With

$$\theta_1^L = 3/2,\ \theta_1^H = 2,\ \theta_2^L = 1/4,\ \theta_2^H = 2,\ c = 1, \tag{1.4}$$

for instance, we can verify that this is true, so that the expected externality mechanism is not ex-post fairness incentive-compatible.

To achieve ex-post fairness implementation, we can instead apply the construction provided in the proof of Proposition 1.1 and augment the AGV by giving player $i = 1, 2$ the option to take an additional amount of $e_{ij} = 1/[(1-\alpha)y_{ji}]$ from player $j \neq i$. With

$$\alpha = 1/2,\ y_{12} = 1,\ y_{21} = 1, \tag{1.5}$$

for instance, we obtain $e_{12} = e_{21} = 2$. Truthful type revelation without claiming this additional payment is then still an IFE-TI, but the associated kindness values are now $\kappa_{ij}(s^*) = 1/y_{ji}$ instead of zero. Updating leaves these values unchanged, so that ex-post utility coincides with the sum of ex-post material payoffs. No player therefore regrets having helped to induce a materially Pareto efficient allocation, or not having taken more money from the other player.

1.5.4 Voluntary Participation

Consider finally the possibility that each player can veto the AGV mechanism on the interim stage, thereby inducing the null allocation $\bar{a} = (0, 0)$ instead of the allocations described in Table 1. Without social preferences, player 2 of type θ_2^L would strictly prefer to do so whenever $(\theta_1^H - \theta_2^H)/2 + (\theta_2^L - c)/4 < 0$, which is again the case for the parameters introduced in (1.4) above. The same holds with intention-based social preferences. It is easily verified that both types of player 1 suffer in material terms from using the veto, provided that player 2 always tells the truth. Since $P_2 > 0$, the veto also hurts player 2's expected material payoff, which makes it an inefficient action for both types of player 1.

Inefficient actions do not influence equitable payoffs, so truth-telling of player 1 remains associated with a type-invariant kindness of zero, from our earlier arguments. Player 2 then cares for material payoffs only, and still prefers to opt out of the mechanism if type θ_2^L has realized.

To guarantee voluntary participation, we can use the construction provided in the proof of Proposition 1.2. With the parameters given in (1.4) and (1.5), case (b) from the proof applies. As before, we obtain the payments $e_{12} = e_{21} = 2$ that each player can claim from the other if none of them makes use of the veto. We obtain the smaller payments $d_{12} = 11/8$ and $d_{21} = 27/16$ that can be claimed when at least one player makes use of the veto. The value of d_{ij} is defined by the equality $-d_{ij} = P_j - e_{ij}$ and ensures that the maximal damage that player i can do to j by using a veto strategy is the same as by using a type-announcement strategy.

1.6 Conclusion

We have studied notions of robustness in implementation for a mechanism design framework where agents are characterized by intention-based social preferences. Within this model, players are uncertain about their opponents' intentions, as they are uncertain about their material payoff types. We have firstly addressed robustness with regard to assumptions about how agents accommodate the uncertainty about others' intentions. Our concept of implementation in IFE-TI provides robustness in this regard, as it renders intentions transparent despite the presence of asymmetric information. We have secondly proposed a notion of ex-post fairness implementation, which provides robustness to the extent that no player would want to change his interim decision even if he were informed about the others' private information ex-post. This concept rules out any ex-post regret.

As our main result, we have established that any materially Pareto-efficient social choice function which provides insurance can be implemented under both robustness concepts, even if participation in the mechanism is voluntary. The insurance property is essential to our construction, because it facilitates the property of type-invariant kindness. The mechanisms that we construct allow the designer to manipulate reference points for kindness perceptions in order to align individual and social motives both on the interim stage and on the ex-post stage, even if participation is voluntary.

Chapter 2

Honesty and Relational Contracts

This chapter shows experimentally how asymmetric information and resulting doubt about cooperation affect the performance of relational contracts. We study the repeated interaction between a principal and her agent where the agent holds private information about his costs of effort provision being either high or low. At the beginning of the interaction, each agent can choose to either truthfully or dishonestly signal his cost type to his principal. We observe that many low cost agents decide to signal high costs. Principals facing high cost signals are uncertain about whether they are facing an honest high cost agent or a dishonest low cost agent who attempts to exploit the informational asymmetry to the own benefit. However, doubt about the truthfulness of high cost signals does not adversely affect the provision of efficiency relative to our control treatment with high cost agents and complete information. This result follows despite the fact that dishonesty leads to substantial payoff inequality and strongly reduces principals' payoffs. Overall, we find that asymmetric information harms the provision of efficiency if dishonest low cost agents imitate the behaviour of high cost agents. We furthermore show that our experimental findings can be organized using a theoretical model of relational contracts.

2.1 Introduction

Many economic relationships utilize relational contracts, informal agreements established and sustained within a relationship, to facilitate cooperation and implement efficient transactions (see e.g. MacLeod (2007) and Malcomson (2012b) for surveys). Informational asymmetries, pervasive in many economic relations, impede the identification of what constitutes cooperative behaviour. In particular, asymmetric information can preclude one party to verify whether the other honours a relational contract or takes an unfair advantage to the own benefit. Suspicions about another's unfair play may lead to allegations, false accusations and breed conflict within a relationship. This suggests that the provision of efficiency through relational contracts suffers in the presence of asymmetric information.

In this chapter, we show experimentally how asymmetric information and doubt about fair play affect the performance of relational contracts. We study the relationship between a principal and her agent, comprised of a sequence of transactions. Within our main treatments, the principal first pays a wage and the agent chooses a costly effort thereafter, to conclude one transaction. A principal's profit increases in effort and her provided wage benefits the agent's payoff. As the parties cannot enter into binding agreements to behave, they must rely on informal agreements to benefit from the relationship. Before the interaction begins, each agent is randomly assigned to either low or high costs of effort provision. In our complete information treatments, each principal is informed about her agent's cost type. Therefore, she can identify the surplus as well as her agent's payoff from a transaction. Put differently, each principal can verify whether her agent acts cooperatively or takes an unfair advantage.

In our incomplete information treatments, each agent is privately informed about his cost type. Prior to the interaction taking place, each agent sends a signal to his principal concerning his cost type. Each principal knows that the assignment to either the low or the high cost type is random, that her agent can choose his signal independent of his actual costs and no principal can verify whether the signal constitutes an honest report or not. Hence, a principal presented with a high cost signal can not be sure whether she is facing an honest high cost or a dishonest low cost agent who exploits the informational asymmetry in pretending to have high costs. This uncertainty can breed doubt about the agent's fair play. For instance, a principal facing a high cost signal cannot be sure whether a transaction constitutes a fair deal with an honest high cost agent or generates a substantial payoff advantage for a dishonest low cost agent.

A large fraction of low cost agents within our main incomplete information treatment chooses to be dishonest: about two thirds of all low cost agents send high cost signals to their principals. Elicited beliefs suggest that such behaviour is common knowledge among the parties. Hence, principals who receive a high cost signal doubt its truthfulness. Yet, our evidence shows that such doubt does not adversely affect the provision of efficiency. Relationships in which the agent provides a high cost signal do not generate significantly different levels of efficiency compared to relationships with high cost agents and complete information, in which doubt is ruled out by design. Moreover, principals who receive a low cost signal are almost certain about its truthfulness. Their agents' choice of honesty thus essentially resolves the informational incompleteness. We observe that efficiency in such relationships is not significantly different from efficiency in case of low costs and complete information.

Overall, asymmetric information reduces efficiency in our main treatments to the extent that relationships with low cost agents generate less efficient outcomes under incomplete relative to complete information. This finding can be organized as follows. On the one hand, honest low cost agents induce no different efficiency than low cost agents under complete information. On the other hand, dishonest low cost agents provide efficiency as if they were assigned to high costs. High costs, however, lead to inferior levels of efficiency compared to low costs in the context of complete information. Because two thirds of all low cost agents choose to lie, efficiency suffers under asymmetric relative to complete information.

The observation that efficiency with high cost signals under incomplete information is not different from efficiency with high costs under complete information contradicts our initial expectation that doubt about the fair play of the agent would destabilize a relationship. This is particularly surprising because such doubt would have been justified: the difference in payoffs between a principal and her agent is about 11CHF (12$) higher in case of dishonest low cost compared to honest high cost agents. Moreover, dishonesty harms the principals' profits: those matched with dishonest low cost agents earn about 15CHF (16$) less than those paired with honest low cost agents. Yet, principals seem to take little action in response as they pay higher wages for a given effort level if they receive a high cost signal compared to a low cost signal. This suggests that doubt does not adversely affect efficiency because principals do not let doubt affect their behaviour.

To investigate the robustness of the above insights, we address the behaviour observed in additional treatments which are identical to our main conditions with the exception of the order of moves within each round of transacting. In particular, each round within our additional treatments takes place in the following order: the agent first exerts effort and the principal provides a wage thereafter. This implies that the principal always has the final say in the distribution of payoffs from a transaction. Our data show that this feature empowers the principals to let their doubt about the truthfulness of high cost signals translate into their behaviour: principals who receive a high cost signal do not pay different wages for a given effort level than principals who face a low cost signal.

Nevertheless, doubt about high cost signals does not adversely affect efficiency in our additional treatments. This corroborates the evidence observed in our main treatments. Yet, in our additional treatments, dishonest low cost agents provide no different levels of efficiency than low cost agents under complete information. In other words, dishonest low cost agents in these treatments act as if they had not lied, unlike such agents in our main treatments who imitate the behaviour of high cost agents. This implies that asymmetric information within our additional treatment does not harm the provision of efficiency compared to complete information.

Our empirical analysis is complemented by a theoretical investigation. We show that the empirically observed patterns can be organized by a theoretical model of relational contracts. In particular, the results of our theoretical comparative statics analysis are consistent with the observed relationship between wages and efforts as well as its dependency on the cost types. For the case of incomplete information, we show that the conditions characterizing pooling equilibria are equivalent to the conditions in case of high costs and complete information. This prediction is consistent with the empirical observation of pooling in effort among dishonest low cost and honest high cost agents in our main incomplete information treatment and the finding that such pooling leads to no different outcomes than under complete information and high costs. Moreover, we show that simple separating equilibria exist. This suggests that subjects within our main treatments choose to pool even though separation could have been feasible.

Related to this study is the experimental literature on relational contracts, especially the work of Brown et al. (2004, 2012).[1] These authors consider principals and agents interacting in an experimental market environment in which the parties themselves determine their matching into pairs as well as the duration of their relationships. The authors show that, in the absence of opportunities to write explicit contracts, long-term bilateral relationships sustained by relational contracts emerge endogenously.[2] Our research builds on this finding. In particular, we adapt the design proposed by Brown et al. (2004, 2012) to the extent that principals and agents within our experiment are randomly assigned into bilateral relationships lasting for a given number of rounds. This simplified design allows us to cleanly identify the effect of asymmetric information and doubt on the performance of relational contracts.

Relational contracts under asymmetric information have been addressed within a small but growing literature (see e.g. Levin (2003), Halac (2012), Englmaier and Segal (2012), Malcomson (2012a,b), Yang (2013) and Li and Matouschek (2014)). However, this research has primarily focussed on theoretical insight. Empirical results are scarce. We therefore contribute to this literature by providing experimental evidence concerning the consequences of asymmetric information on the performance of relational contracts.

Our research is more broadly related to the literature on honesty and deception (see e.g. Gneezy (2005), Sanchez-Pages and Vorsatz (2007), Mazar et al. (2008), Hurkens and Kartik (2009), Pruckner and Sausgruber (2013), Gibson et al. (2013) and Fischbacher and Foellmi-Heusi (2013)). This research has discussed the measurement as well as the potential motives for dishonesty. We also document dishonest behaviour in our experiment. Yet, the focus of our research lies on the consequences of dishonesty on efficiency.

[1] See e.g. Wu and Roe (2007), Fehr et al. (2009) and Camerer and Linardi (2010) for studies building on the research of Brown et al. (2004, 2012).

[2] The evidence provided by Brown et al. (2012) illustrates that relationships which last for multiple periods emerge irrespective of whether there is excess demand or excess supply for agents' labour in a market. However, these authors also report that more long-term relationships emerge in markets in which the number of agents exceeds the number of principals compared to markets in which there are more principals than agents.

The remainder of the chapter is organized as follows. We present the design of our experimental study in Section 2. We state our main behavioural predictions within Section 3 and report the empirical results of our study in Section 4. We then present theoretical considerations with regard to our empirical findings in Section 5. Section 6 concludes.

2.2 Experimental Design

2.2.1 Treatments

As noted in the Introduction, our design builds on the research of Brown et al. (2004, 2012) and focuses on the bilateral relationships between principal and agents. As our main treatment variable, we vary the principal's information about the agent's payoff type. Prior to the interaction taking place, each agent's type is randomly determined and communicated to the agent. In our treatments with complete information, a principal also gets informed about her agent's actual type before the interaction begins. Within our conditions featuring incomplete information, the agent holds private information about his type. A principal is informed that her agent's type is randomly determined and that she cannot obtain definite information about her agent's true type throughout the experiment. Our incomplete information conditions additionally feature signalling: after learning the type, each agent sends a message concerning his type to his principal. The agent can select the signal independent of his true type and this is known to the principal.

We devised two main treatments, PC and PI, where PC features complete and PI incomplete information with signalling prior to the interaction, as described above. A relationship is comprised of a sequence of transactions each of which takes place sequentially. In the context of the conditions PC and PI, the principal moves first and the agent thereafter at each transaction. In particular, a principal first chooses a wage and indicates a desired effort level. The principal's wage payment is binding. An agent observes his principal's choice and selects an effort level thereafter. Importantly, an agent is free to choose his effort and does not have to follow the desired effort level indicated by his principal. In other words, effort can not be enforced within the treatments PC and PI.

We designed two additional treatments AC and AI where condition AC is characterized by complete and condition AI by incomplete information with signalling as outlined the first paragraph of this section. The difference between these treatments and our main conditions PC and PI concerns the order of moves in which the parties conduct a transaction. In the context of AC and AI, the agent moves first and the principal thereafter in each round. In particular, an agent first chooses an effort and indicates a desired wage. The agent's choice of effort in these conditions is binding. After observing the agent's choices, the principal chooses a wage. The desired wage indicated by the agent is non-binding for the principals' choice and, therefore, wages can not be enforced in the context of the conditions AC and AI.

Table 2.1: Treatments

Treatment	First-Mover	Information
PC	Principal	Complete
PI	Principal	Incomplete
AC	Agent	Complete
AI	Agent	Incomplete

2.2.2 Parameters

In the context of all treatments, a relationship consists of a sequence of 15 transactions. In each of these, a principal chooses a wage w from $[0, 100]$ and an agent selects an effort e from $\{1, 2, \ldots, 9, 10\}$. A desired wage w^d and desired effort e^d are chosen from the same sets as wages and efforts are selected, respectively. For given choices of w and e within a round, a principal's material payoff $\Pi_P(w, e)$ and her agent's material payoff $\Pi_A(w, e, \theta)$ are given by

$$\Pi_P(w, e) = 10 \cdot e - w \text{ and } \Pi_A(w, e, \theta) = w - c(e, \theta)$$

where $c(e, \theta)$ indicates the cost of effort e given cost type $\theta \in \{L, H\}$ as summarized in Table 2.2 below. Costs are strictly increasing and marginal costs are weakly increasing in effort, irrespective of the agent's cost type. For any given $e > 1$, low costs $c(e, L)$ are strictly lower than high costs $c(e, H)$. The difference between high and low costs, $c(e, H) - c(e, L)$, is increasing in e. Yet, since the marginal benefit of effort to a principal always strictly exceeds the marginal cost of effort to her agent, the efficient effort level equals the highest feasible effort under both cost regimes.

Table 2.2: Cost of Effort

e	1	2	3	4	5	6	7	8	9	10
$c(e, L)$	0	0.5	1	2	4	6	8	10	13	16
$c(e, H)$	0	3	6	10	15	20	25	30	36	42

In all experimental conditions, and before the interaction took place, every agent is randomly assigned to either type L or type H and then privately informed about it. After observing their type, agents within the incomplete information conditions choose between the message "I have low costs." and the message "I have high costs.", irrespective of their actual type. Hence, agents are free to either be honest or to lie about their type. A message can only chosen at the beginning of the experiment and can not be reversed in the course of it. A principal receives the message selected by her agent and is informed that she will not be able to obtain definite information about her agent's true costs in the context of the experiment. Within the complete information conditions, a principal receives either message "Your seller has high costs." or message "Your seller has low costs." depending on the true costs of her agent.

2.2.3 Procedures

The experiment is framed in the neutral language of a goods exchange where a principal is referred to as a "buyer" and the agent as a "seller".[3] Assignment into the roles of a buyer and seller is random, the matching of principals and agent is randomly determined, and each match persists over fifteen rounds of transacting. At the end of each period, both principal and agent receive a summary of their choices in the current round including the wage and effort as well as the desired wage and desired effort, depending on the treatment. Every player is additionally informed about the own material payoff in the current round which is noted in terms of the experimental currency "Punkte" (points). The sum of payoffs, taken over all rounds, is converted into real money by the end of the experiment (10 Points=1 CHF($1.1)) and paid out in combination with the show up fee (100 points). In the context of the incomplete information conditions, we additionally elicited principals' first-order beliefs about the (dis-)honesty of agents after the interaction had taken place.[4] We furthermore elicited agents' second order beliefs about their principals' first order beliefs.[5] All elicited beliefs are incentivized.[6]

Table 2.3: Number of Participants

Treatment	Sessions	Subjects per Treatment
PC	2	56
PI	4	128
AC	2	64
AI	2	64

The experiment was computerized using the software z-tree (Fischbacher, 2007). For organizing and recruitment, we used the software hroot (Bock et al., 2012). Our subject pool consists primarily of students at the University of Zurich and the Swiss Federal Institute of Technology in Zurich. In total, 312 subjects participated in the experiment in August, October and November of 2013 and April of 2014. No subject participated in more than one session. On average, a session lasted 95 minutes with an average payment of 48 CHF ($52).

[3]See Appendix A, Section A.4 for sample instructions.

[4]We elicited beliefs about all other agents' message choices, rather than the belief about a principal's actual agent. For instance, in a session with 16 agents where 8 are assigned to low costs, we asked the following question: "8 out of 16 sellers were assigned to low costs. How many of these sellers with true low costs sent the message "I have low costs." to their buyers?".

[5]In a session with 16 agents where 8 are assigned to low costs, agents answered the following question: "Your buyer was asked the following question: '8 out of 16 sellers were assigned to low costs. How many of these sellers with true low costs sent the message "I have low costs." to their buyers?' What do you believe: which answer did your buyer provide in response to this question?".

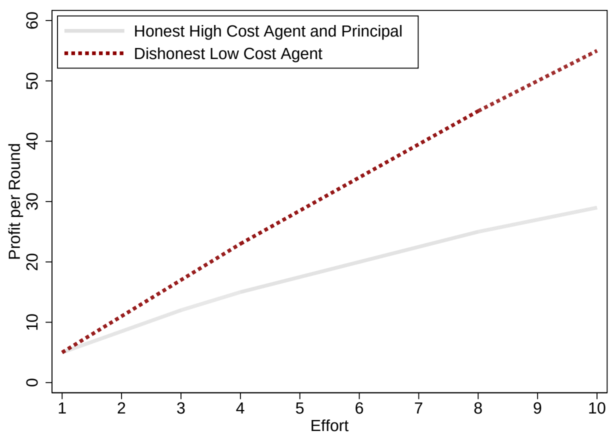
[6]Subjects earn, on average, about 20 additional points for indicating their beliefs.

2.3 Predictions

Principals and agents within the experimental environment outlined above cannot enter into binding agreements. The parties may, however, utilize relational contracts to facilitate efficient transactions within their relationships. For an illustration, suppose a principal and her agent implicitly agree to exert efficient effort in exchange for a fair wage, i.e. a payment which yields equal profits for both parties from the transaction. Such behaviour can be sustained if each party believes that deviating from the agreement to one's own short-term benefit will be punished in the course of the relationship and therefore not pay out.[7] Because a principal and her agent can always compute each others' profits from a given transaction in the complete information conditions, they can always verify if the other honours the agreement or takes an unfair advantage to the own benefit.

By contrast, doubt about whether the other party plays fair may arise in the presence of asymmetric information. Suppose the principal receives a high cost signal from her agent. Hence, the principal cannot not be sure whether she is facing an honest high cost or a dishonest low cost agent. The principal therefore cannot pin down her agent's actual payoff. But nevertheless she can determine her agent's potential profits. For an illustration, suppose a transaction constitutes a fair deal between a principal and an honest high cost agent, i.e. both parties obtain equal profits. As Figure 2.1 indicates, such a transaction would imply that a dishonest low cost agent earns up to twice as much profit as her principal. Hence, if a principal receives a high cost signal, then she may not be sure if a given transaction constitutes a fair deal or comes with substantial payoff inequality.

Figure 2.1: Profits from a Fair Transaction in Case of High Costs



Notes: For every given effort, the wage inducing equal profits between a principal and a high cost agent was computed. If a principal provides such fair wages, then her profits as well as an honest high cost agent's payoffs are described by the grey solid line whereas a dishonest low cost agent's profits are illustrated by the red dashed line, where linear interpolation was applied.

Doubt about an agent's honesty and his fair play may breed conflict in a relationship and adversely affect its performance. For instance, a principal who suspects a high cost

[7]See Section 2.5 for formal descriptions of behaviour prescribed by relational contracts.

signal to constitute a dishonest report may no longer be willing to pay wages which would be fair in case of true high costs. However, the provision of less than such wages may be considered unfair by an honest high cost agent and therefore be punished with reductions in effort which, in turn, implies lower efficiency. A dishonest low cost agent may imitate such behaviour in order to keep pretending to be assigned to high costs. As a consequence, efficiency provided in the context of incomplete information and high cost signals could be lower than efficiency under complete information and high costs, in which case such doubt is ruled out by design. However, it is not obvious whether this prediction prevails. A principal may not let doubt about the truthfulness of a high cost signal and her agent's fair play translate into her behaviour. This may guard the relationship from additional conflict and therefore prevent that efficiency is lower than in case of complete information and high costs. Yet, as illustrated above, such argument rests on the assumption that a principal is prepared to live with the possibility of being treated unfairly by a dishonest agent.

In light of the above line of argument, it seems questionable whether low cost agents select high cost signals in the first place. Indeed, a truthful report of low costs may virtually eliminate the informational asymmetry, since there seem to be no obvious reasons for high cost agents to send low cost signals. This may pave the way for a relationship no less efficient than under complete information and low costs. Moreover, a truthful report of low costs and therefore the explicit choice to refrain from attempting to take an unfair advantage may indicate the willingness to behave fairly to a principal. This may foster a relationship from the start and therefore even enhance its performance beyond the level provided under complete information.

To sum up, the following questions defy a clear theoretical answer, suggesting a need for careful experimental analysis. First, do low cost agents select high cost signals or do they report truthfully? Second, is efficiency lower with high cost signals and incomplete information than with high costs and complete information? The answer to this question sheds light on the effect of doubt on the performance of relational contracts. Third, is efficiency higher with low cost signals and incomplete information than with low costs and complete information? This question concerns the consequences of honesty on relational contracts. Fourth, is efficiency lower with low cost types and incomplete information compared to low costs and complete information? The answer to this question allows us to judge the effect of asymmetric information on the functioning of relational contracts.

2.4 Results

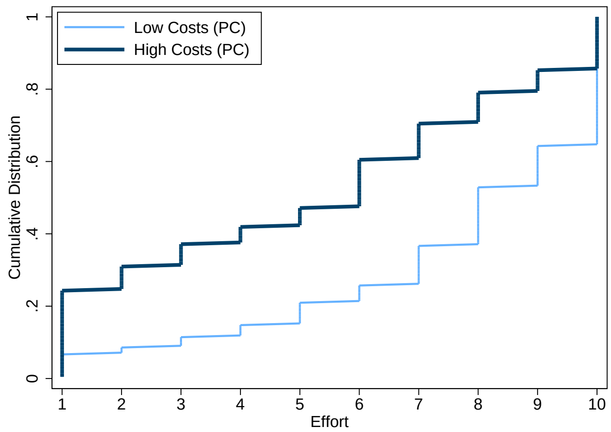
2.4.1 Efficiency and Signal Choice

To lay the ground, we address the case of complete information and report on the consequences of the different cost types on efficiency. We are not aware of any previous experimental evidence in the regard. Documenting such an effect thus closes a gap concerning the empirical understanding of relational contracts. Moreover, it provides a basis for understanding the relationship between efficiency and (dis-)honesty which directly concerns the cost types. As the data elicited within our main complete information treatment PC show, cost types affect the provision of efficiency.

Result 2.1. *Effort in case of low costs is higher compared to high costs in treatment PC. Hence, cost types affect efficiency under complete information.*

Average effort provided by low cost agents in condition PC equals 7.58. By contrast, average effort in PC exerted by high cost agents equals 5.23. The difference in average effort is significant (p-value 0.01, t-test[8]). Figure 2.2 presents the cumulative distributions of effort, which further illustrate the differences in effort depending on the cost regime. As we argue in Section 2.5.1, this empirical finding can be organized within a theoretical model of relational contracts by comparing the equilibrium sets in cases of high and low costs.

Figure 2.2: Cost Types and Effort in PC



The performance differences between relationships with high and low cost agents raise the question whether agents within the incomplete information treatments decided to truthfully signal their types or not. In light of the above results, one may expect low cost agents to truthfully signal their type, hoping to benefit from a highly productive relationship. On the other hand, low cost agents may report high signals in the hope for higher wages and thus higher payoffs. However, as noted in Section 2.1, high cost signals may breed doubt about whether an agent plays fair and therefore lead to potentially less productive

[8] In the following, all t-tests concerning effort are two-sided and feature clustering by the individual.

relationships. The observed empirical data show that a large fraction of low cost agents choose dishonesty.

Result 2.2. *(i) Two thirds of all low cost agents selected a high cost signal in PI.*
(ii) Elicited beliefs in this treatment suggest that principals were aware of the extent of dishonesty and that their agents suspected so.

We find that about 66% of all low cost agents participating in condition PI chose a high cost signal. True high cost agents, by contrast, chose to be honest: all high cost agents - with the exception of one agent - reported truthfully. Principals' first-order beliefs indicate that such behaviour was roughly expected. On average, they believed about 71% of low cost and about 6% of high cost agents to be dishonest. Principals thus even slightly overestimated dishonest behaviour. Agents' second-order beliefs about their principals' first-order beliefs suggest that agents were well aware of their principals' suspicions concerning dishonesty. On average, agents believed that their principals believed that about 63% of low cost and 9% of high cost agents reported untruthfully.

Table 2.4: Dishonesty among Low Cost Agents in PI

Observed Dishonesty	66 %
1^{st} order beliefs	71 %
2^{nd} order beliefs	63 %

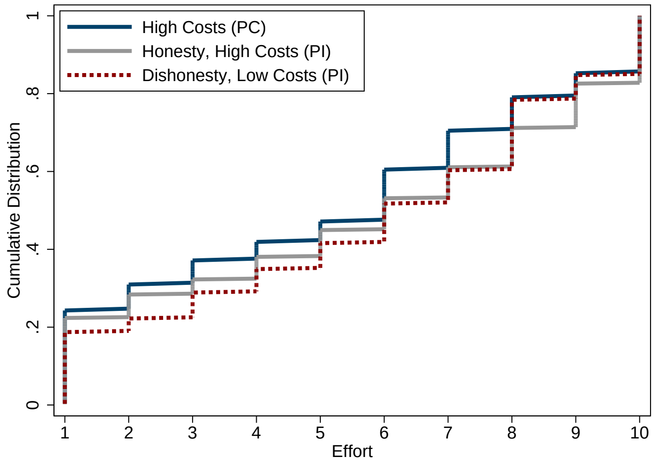
Since principals believed that a substantial fraction of low cost agents provided a dishonest signal, they must have been aware that either a truthful high cost or a dishonest low cost agent could be hiding behind a high cost signal. As discussed in Section 2.3, the consequences of doubt about the truthfulness of high cost signals were not obvious ex ante. On the one hand, one might conjecture that doubt adversely affects the performance of relationships to the extent that efficiency under high cost signals is lower than under high costs and complete information, where doubt is absent by construction. On the other hand, it is possible that doubt does not translate into behaviour if principals accept the possibility of deception and payoff inequality. In this case, efficiency under high cost signals in PI may be no different from efficiency under high costs in PC. Indeed, this prediction is borne out by the data.

Result 2.3. *Effort under high cost signals in PI does not differ from effort under high costs in condition PC. In this sense, doubt does not adversely affect efficiency.*

Within treatment PI, average effort exerted by dishonest low cost agents equals 5.78 and does not significantly differ from the average effort provided by truthful high cost agents, which equals 5.66 (p-value 0.86, t-test). Such pooling is further illustrated in Figure 2.3 which presents the distributions of effort. We furthermore find that average effort generated under pooling in case of high cost signals is not significantly different from effort under high costs in the context of condition PC (p-value 0.54, t-test). As we note in Section 2.5.2, these findings can very well be organized within a theoretical model of relational contracts.

Taken together, a high cost signal under condition PI has no different consequences for efficiency than high costs in condition PC, despite the fact that parties were suspecting a substantial degree of dishonesty.

Figure 2.3: Effort in Cases of High Cost Information



Having addressed the consequences of doubt on efficiency, we now turn to the case of low cost agents who chose to be honest. Because principals believed low cost signals to be truthful with high probability, an honest report of low costs should have nearly resolved the informational asymmetry. As noted in Section 2.3, this may imply that relationships informed by low cost signals perform no different from relationships with low costs and complete information. However, because both principals and agents believed that a large fraction of low cost agents behaved dishonestly, the choice of a low cost signal, and therefore the explicit choice not to attempt to take an unfair advantage, may have nourished the performance of such relationships. However, as the following result shows, this prediction was not confirmed by our data.

Result 2.4. *Effort provided by honest low cost agents in condition PI does not differ from effort exerted by low cost agents in treatment PC. In this sense, honesty does not enhance efficiency.*

On average, average effort among truthful low cost cases within condition PI equals 7.3 and is not significantly different from average effort provided by low cost agents in treatment PC (p-value 0.77, t-test).[9] The evidence presented so far therefore suggests that asymmetric information reduces the provision of efficiency by low cost agents. On the one hand, the explicit choice of honesty does not further efficiency beyond the level provided under complete information. On the other hand, the choice of dishonesty reduces efficiency since dishonest low cost agents imitate the behaviour of high cost agents and therefore provide inferior levels of effort than low cost agents in our complete information treatment PC. Given that a substantial fraction of low cost agents chose dishonesty in PI,

[9]The distributions of effort, as presented in Appendix A.1, Figure A.1, further illustrate this result.

we therefore find that asymmetric information harms the provision of efficiency: average efficiency among low cost agents in condition PI equals 6.31 and is lower than average efficiency provided by low cost agents in the context of condition PC (p-value 0.07, t-test).[10]

Result 2.5. *Effort provided by low cost agents in PI is lower than effort selected by low cost agents in PC. Hence, asymmetric information reduces efficiency under low costs in our main treatments.*

2.4.2 Profits and Wages

As we argued in Section 2.3, there are reasons to believe that doubts can potentially lead to adverse consequences for the performance of relational contracts. Our prediction in this regard rested on the joint hypothesis that (a) dishonesty goes along with advantageous inequality for the dishonest low cost agent and (b) the principal reacts adversely to the possibility of such inequality. We now argue that the observed absence of efficiency consequences reflects the failure of the second rather than the first hypothesis. Our evidence shows that dishonesty leads to payoff advantages for untruthful agents and imposes profit reductions on their principals.

Result 2.6. *(i) The difference in profits between a principal and her agent is about 11CHF (12$) higher in case of a dishonest low cost agent compared to an honest high cost agent.*
(ii) Dishonesty does not raise low cost agents' profits compared to honesty.
(iii) Principals earn about 15CHF (16$) less if they are paired with dishonest low cost compared to an honest low cost agents.

The average difference between dishonest low cost agents' and their principals' profits amounts to 192 points in PI. This average is significantly different from the average difference between honest high cost agents' and their principals' profits, which equals 85 points in treatment PI (p-value 0.01, Wilcoxon rank-sum test).[11] In other words, if the principal receives a high cost signal then the difference between her and her agent's profits is about 11CHF (12$) higher if she is facing a dishonest low cost agent than if she is matched with an honest high cost agent. This finding strongly suggests that the absence of an effect of doubt on efficiency does not reflect the absence of advantageous inequality for the dishonest low cost agent.

Moreover, we find that dishonesty does not pay for low cost agents. On average, honest low cost agents earn 503 points while dishonest low cost agents make 478 points in PI (p-value 0.74, Wilcoxon rank-sum test). While the choice of dishonesty rather than honesty does not enhance low cost agents' own profits, it substantially harms their principals'. In particular, principals matched with honest low cost agents earn, on average, 440 points whereas average earnings of those paired with dishonest low cost agents equal 286 points

[10]See Appendix A.1, Figure A.2, for the distributions of effort.

[11]In the following, the indicated points exclude the show up fee as well as the points earned for beliefs and thus exclusively reflect subjects' earnings within their relationships.

in PI (p-value 0.01, Wilcoxon rank-sum test). Put differently, by choosing to be untruthful low cost agents deprive their principals by about 15 CHF (16$).

Having established that principals doubt high cost signals and that dishonesty generates inequality and payoff reductions, we now investigate whether this affects the principals' behaviour. In particular, we address how information about the cost type affects the provision of wages. On the one hand, doubt about the truthfulness of high cost signals could translate into principals' behaviour. In order to prevent dishonest low cost agents from taking an unfair advantage, principals could pay the same wages for given effort if they receive a high cost signal compared to if they get a low cost signal. This however, may lead to conflict with honest high cost agents who could consider such wages unfair. Hence, on the other hand, principals could condition their wages for a given effort level on the indicated costs, despite their doubt about high signals, to prevent dispute within the relationship. This prediction is consistent with our evidence.

Result 2.7. *(i) Wage and effort are positively related.*
(ii) Wage payments for a given effort level depend positively on the indicated costs.

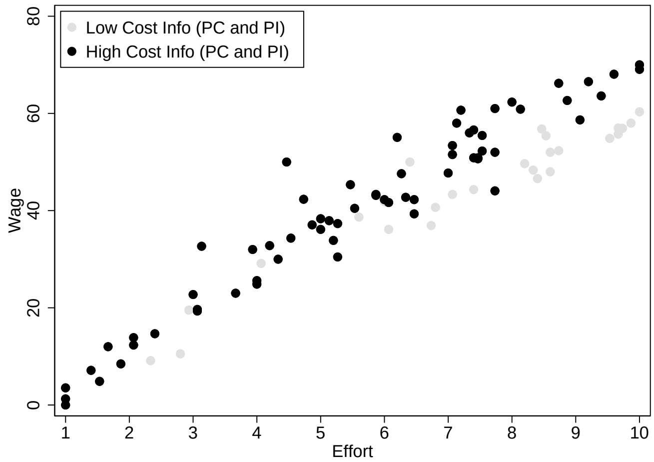
We find a strong positive relation between wages and effort within our treatments PC and PI. This observation is in line with prior evidence on relational contracts in the laboratory, for instance as presented by Brown et al. (2004). Within our treatments PC and PI, an additional unit of effort significantly increases average wages by about 6.6 points as shown by a regression analysis.[12] Moreover, wages for a given effort level are significantly higher if agents are assigned to high relative to low costs in PC. [13] In addition, principals in condition PI pay higher wages if they receive a high compared to a low cost signal.[14] This shows that principals' doubt about high cost signals did not translate into their wage setting behaviour. Figure 2.4 illustrates the relationship between wages and efforts as well as its dependency on the indicated costs. As we show in Section 2.5, these empirical findings are consistent with the corresponding comparative statics results in a theoretical model of relational contracts.

2.4.3 Robustness

The evidence presented so far suggests that doubt does not adversely affect efficiency, despite the fact that dishonesty increases inequality and harms principals' profits, because principals still pay higher wages in case of high compared to low cost signals and therefore avoid introducing conflict into a relationship. In the following, we investigate whether the observed effects are specific to the design of our treatments PC and PI. In particular, every transaction within these conditions was conducted in the same order: the principal moved first and the agent thereafter. Hence, an agent always had the final say in the distribution of payoffs within a round. Our evidence shows that being in the position of a

[12]See Appendix A.1, Table A.1, Column (1).
[13]See Appendix A.1, Table A.1, Column (2).
[14]See Appendix A.1, Table A.1, Column (3).

Figure 2.4: Wage-Effort Relationship



Notes: One data point in the above plot represents the vector of average wage and average effort chosen in one relationship between a principal and an agent, where the averages are taken over all periods. All data points were elicited within condition PC and PI. "Low Cost Info" ("High Cost Info") represents relationships in which either principals received low (high) cost signals in the context of PI or costs were low (high) within PC.

second-mover empowered agents relative to their principals: agents' average profits in PC and PI are about 117 points higher than their principals' profits (p-value 0.00, t-test[15]). As a consequence, acting in the comparatively less influential position of a first-mover may have rendered it infeasible for the principals to challenge high cost signals, meaning that they pay no higher wages in response to high compared to low cost signals. Therefore, the order of moves within our treatments PC and PI may have impeded doubts to affect efficiency.

To investigate this claim, we address behaviour observed in our conditions AC and AI. Within these treatments, the agent moved first and the principal thereafter in order to conduct a transaction. Hence, in these treatments, principals could directly shape the distribution of payoffs within each round. Our evidence shows that acting as second-movers put principals' in a more powerful position: their average profits in AC and AI are about 150 points higher than agents' average payoffs (p-value 0.00, t-test). One might therefore conjecture that, empowered by their position as a second-mover, principals' doubt about the truthfulness of high cost signals translates into their behaviour.

A large fraction of low cost agents signalled high costs and principals doubted the truthfulness of such signals in treatment AI.[16] Contrary to condition PI, principals' doubt about high cost signals translated into their behaviour in condition AI: wages for a given effort level in this condition are not significantly higher if principals receive a high compared to a low cost signal.[17] Moreover, the wages provided by almost all principals in AI who received

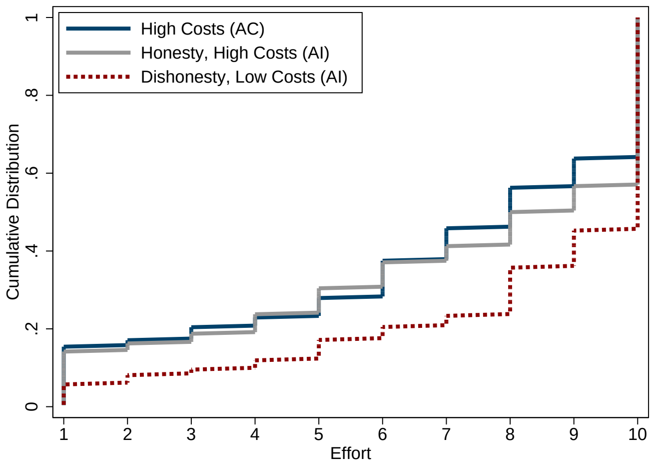
[15]In the following, t-tests concerning differences in profits are two-sided and employ clustering by each principal-agent pair.
[16]About 87 % of all low cost agents in AI chose a high cost signal. Principals believed about 78% of all low cost agents to be dishonest in AI.
[17]See Appendix A.1, Table A.1, Column (4). As Column (5) in this table shows, wages for a given effort

a high cost signal range below the wages which would be fair in case of true high costs.[18] But nevertheless, such behaviour did not adversely affect the performance of relational contracts.

Result 2.8. *(i) Effort under high cost signals in AI does not differ from effort under high costs in AC. In this sense, doubt does not adversely affect efficiency in our additional treatments.*
(ii) Effort provided by low cost agents in AI does not differ from effort selected by low cost agents in AC. Hence, asymmetric information does not harm efficiency for the case of low costs in our additional treatments.

Average effort exerted by agents who sent a high cost signal in AI equals 7.6 and does not significantly differ from average effort chosen by high cost agents in AC which equals 6.9 (p-value 0.31, t-test[19]). This finding is further illustrated by the distributions of effort, as presented in Figure 2.5.[20] Hence, despite the fact that principals do not reward high cost signals with higher wages than low cost signals, we reject the hypothesis that this has negative consequences for efficiency. This shows the robustness of our finding presented in Result 2.3.

Figure 2.5: Effort in Cases of High Cost Information (AC and AI)



However, in contrast to the evidence presented in Result 2.5, low cost agents' average effort in AI equals 8.1 which is not significantly different from low cost agent's average effort in AC, which equals 7.9 (p-value 0.71, t-test). This follows because almost all low

within the corresponding complete information treatment AC are higher if costs are high rather than low.

[18]See Appendix A.1, Figure A.3, for a plot of the wage-effort relationship which emerged under high cost signals in AI.

[19]In the following, t-tests concerning efforts are two-sided and employ clustering on the individual level.

[20]The distributions of effort presented in Figure 2.5 suggest that dishonest low cost and honest high cost agents separate in effort, in contrast to the pooling reported in the context of Result 2.3. This intuition is descriptively confirmed, yet the difference in average efforts turns out to be statistically insignificant: average effort provided by dishonest low cost agents equals 8.23 whereas average effort of honest high cost agents amounts to 7.11 (p-value 0.13, t-test). As our theoretical analysis presented in Appendix A.3.3 shows, both pooling and separation constitute theoretical possibilities.

costs agents in AI choose to lie, but their average effort does not statistically differ from the average effort provided by low cost agents in AC (p-value 0.63, t-test). In other words, dishonest low cost agents in AI behave as if they had not lied. This stands in contrast to the observation that dishonest low cost agents in PI imitate the behaviour of high cost agents in PC, as noted in in the context of Result 2.3. Taken together, asymmetric information, dishonesty and doubt in condition AI do not harm the performance of relational contracts relative to treatment AC.

2.5 Theoretical Considerations

In the following, we interpret our empirical findings using the established theoretical framework for the analysis of relational contracts. We consider the following infinitely *repeated principal-agent (RPA) game*. In each of the periods $t = 0, 1, 2, ...$ a principal and an agent play a two-stage game. The discount factor is $\delta \in (0, 1)$.[21] We suppose that, in the first stage, principals choose wages w_t from $[0, 100]$; in the second stage, agents choose efforts e_t from $\{1, ..., 10\}$. Agents have a cost function $c(e_t, \theta)$, where $\theta \in \{H, L\}$. The function c is increasing in e_t for both cost types, and it satisfies $c(1, H) = c(1, L) = 0$, $c(e_t + 1, H) - c(e_t, H) > c(e_t + 1, L) - c(e_t, L)$ for $e_t \in \{0, ..., 9\}$. The payoff of the principal is $10e_t - w_t$. The payoff for an agent of type θ is $w_t - c(e_t, \theta)$. In Section 2.5.1, we analyse this game under complete information about the cost type. In Section 2.5.2, we turn to the case of incomplete information. This corresponds to an analysis of the interaction starting after the agent has sent a high cost signal. In this case, the common prior of the event that $\theta = L$ is $\mu \in (0, 1)$. Appendix A.2 contains our main formal definitions and proofs; further technical details are provided in Appendix A.3.[22] In the main text, we confine ourselves to an intuitive presentation of the arguments.

2.5.1 Complete Information

Following the literature on relational contracts, we analyse *trigger-strategy equilibria* (TE). These are subgame perfect equilibria in which the players use *trigger-strategy profiles* (TP) described as follows. The principal starts with some $w^* > 0$, and the agent responds with some $e^* > 1$. As long as the history of play consists only of wage choices w^* and effort choices e^*, the principal continues to choose w^* and the agent continues to choose e^*; after a deviation, principals (agents) choose minimal wages (efforts) forever. In addition, we consider *forgiving cut-off strategy equilibria* (FCE). These are subgame perfect equilibria

[21] For reasons of tractability, we designed our experiment as a finite rather than an infinitely repeated game. Figures A.5 and A.6 in Appendix A.1 show that end effects occur primarily in the last period of interaction. Omitting the data of the last round would not qualitatively change the empirical results presented above. Hence, we believe that the theoretical insight gained within an infinitely repeated game framework can help us interpret our data.

[22] Within Appendix A.3 we also present an analysis of the infinitely repeated agent-principal (RAP) game which is identical to the RPA game with the exception of the order of moves within each stage game: within the RAP game, the agent moves first and the principal thereafter in every round. The predictions within this game therefore correspond to the data observed within our treatments AC and AI.

in which the players use *forgiving cut-off strategy profiles* (FCP) such that the principal (agent) sticks to w^* (e^*) unless the other player has initiated a downward deviation from which she has not returned. While the TP is particularly simple to analyze, the FCP has two advantages. Firstly, it appears more plausible to punish only after downward deviations and not after upward deviations. Secondly, punishment behaviour in our data corresponds more closely to forgiving than to trigger strategies.[23]

We shall say that a wage-effort vector (w^*, e^*) is *sustainable as a TE (FCE) for δ and θ* if, for these parameter values, there exists a TE (FCE) such that the resulting outcome path is the infinite repetition of wages w^* and efforts e^*. In this case, we also say that an effort e^* is *sustainable as a TE (FCE) given (δ, θ, w^*)*.

As stated in the following proposition, the conditions required to sustain any given wage-effort vector (w^*, e^*) under complete information are identical for TE and FCE. In this proposition, $S(e^*) = 10e^* - 10$ denotes the additional benefit for the principal from choosing e^* for a fixed w^* rather than the minimal effort 1, for which $S(1) = 0$.

Proposition 2.1. *Consider the RPA game with complete information and cost type $\theta \in \{L, H\}$. Then, a wage-effort vector (w^*, e^*) is sustainable as a TE for δ and θ if and only if*

$$S(e^*) \geq w^*, \tag{2.1}$$

$$w^* \geq \frac{c(e^*, \theta)}{\delta}. \tag{2.2}$$

Moreover, a wage-effort vector (w^*, e^*) is sustainable as an FCE for δ and θ if and only if (2.1) and (2.2) hold.

Proof of Proposition 2.1. See Appendix A.2.1. □

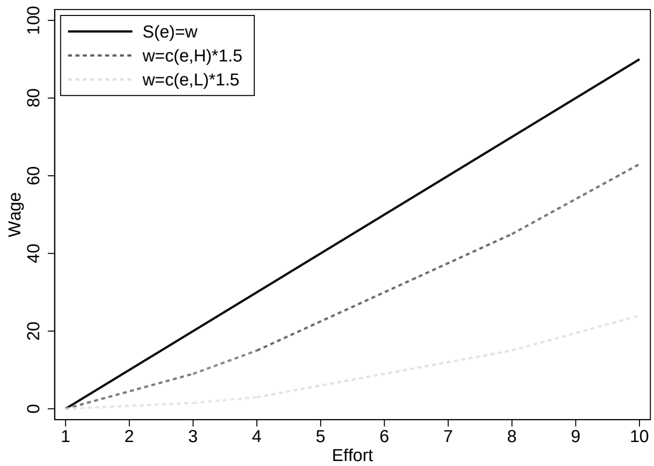
The intuition for the result is similar for TE and FCE. Condition (2.1) guarantees that the principal is willing to pay the wage w^* rather than the minimal wage: She must expect an equilibrium effort of at least e^*, generating benefits $S(e^*)$ that are high enough to compensate for the wages w^*. Condition (2.2) guarantees that the agent who is expecting a normalized discounted equilibrium payoff δw^* is willing to incur the equilibrium effort costs of $c(e^*, \theta)$ rather than exerting the minimal effort.

Proposition 2.1 has several implications that are useful to interpret our empirical observations. For $\theta \in \{H, L\}$, all efforts in $\{1, ..., 10\}$ are sustainable for suitable wages and sufficiently high discount factors: This follows from (2.1) and (2.2) because $S(e) > c(e, H)$ for all $e \in \{1, ..., 10\}$. Figure 2.6 illustrates the sets of sustainable wage-effort vectors for a given parameter choice.[24] The following corollary is immediately intuitive from the figure.

[23]Detailed descriptions of our data are available on request.
[24]We fixed $\delta = 2/3$; similar qualitative patterns emerge for sufficiently large δ.

Corollary 2.1. *(i) The set of e^* that are sustainable as a TE (FCE) given δ and θ is increasing in w^*.*[25]
(ii) The maximal e^* that is sustainable as a TE (FCE) given (δ, θ, w^*) is higher for $\theta = L$ than for $\theta = H$.
(iii) The minimal w^* for which e^* is sustainable as a TE (FCE) given (δ, θ, w^*) is higher for $\theta = H$ than for $\theta = L$.

Figure 2.6: Sustainable Wages and Efforts for $\delta = \frac{2}{3}$



Note: The specifications of c(e,L) and c(e,H) constitute linear interpolations of the low and high cost specifications employed in our experiments, see Table 2.2.

Statement (i) is consistent with the empirical observation that high efforts and high wages go hand in hand (Result 2.7). (ii) states that, if costs increase, higher efforts are harder to sustain with a given wage. This follows because the right boundary of the region described by (2.1) and (2.2) shifts to the left. This provides a possible rationalization of Result 2.1. Similarly, (iii) says that sustaining a higher effort requires higher wages with high costs: the lower bound of the region described by (2.1) and (2.2) shifts up. This statement is in line with the observation that wages conditional on efforts are higher for high costs (Result 2.7).

2.5.2 Incomplete Information

We now proceed to the incomplete information case to understand behaviour in situations in which the agent signals high costs. As reported in Result 2.3, the behaviour of truthful high types and low types who sent the high signal within condition PI under incomplete information is similar. We therefore first search for conditions under which simple pooling equilibria exist. Thereafter, we ask whether separating equilibria exist as well.

[25]The formulation that a set is increasing in some parameter refers to the strong set order; thus both the minimum and the maximum of the set increase.

Pooling Equilibria

We define a *pooling trigger strategy profile* (PTP) by the requirement that principals and agents play a trigger strategy as described in Section 2.5.1 and both types of agents choose the same strategy (that is, e^* is the same for both types). The definition of the *pooling forgiving cut-off strategy profile* (PFCP) is analogous. A *pooling trigger strategy equilibrium (PTE)* is a weakly perfect Bayesian equilibrium in which the players choose a PTP. A *pooling forgiving cut-off strategy equilibrium (PFCE)* is a weakly perfect Bayesian equilibrium in which the players use a PFCP. The adaptation of the sustainability definition to pooling equilibria is straightforward.

Proposition 2.2. *Consider the RPA game with incomplete information. Then, a wage-effort vector* (w^*, e^*) *is sustainable as a PTE given* δ *if and only if*

$$S(e^*) \geq w^*, \tag{2.3}$$

$$w^* \geq \frac{c(e^*, H)}{\delta}. \tag{2.4}$$

Moreover, a wage-effort vector (w^*, e^*) *is sustainable as an PFCE for* δ *if and only if (2.3) and (2.4) hold.*

Proof of Proposition 2.2. See Appendix A.2.2. □

Thus, the condition for sustainability with incomplete information exactly corresponds to the condition for sustainability under high costs with complete information. Intuitively, in a pooling equilibrium the binding incentive constraint is that the high-cost types remain on board and are willing to exert effort. Proposition 2.2 is consistent with Result 2.3: behaviour of high-cost agents in the incomplete information treatment following high cost signals is very similar to behaviour in the complete information condition with high costs. Moreover, Proposition 2.2 obviously implies that, for fixed discount factor δ, the set of sustainable e^* is increasing in w^*. This statement, which is the incomplete information analogue of the first part of Corollary 2.1, is consistent with the observed positive relation between wages and efforts in condition PI (Result 2.7).

Separating equilibria

In the following, we show that a particularly simple type of separating equilibrium exists for a large set of parameters and efforts. In a *separating trigger-strategy equilibrium (STE) sustaining* e^H *and* e^L, the principal chooses a wage $w^* \geq 0$ in the first period. Then high and low cost types separate by choosing efforts e^H and e^L, respectively, where $e^H < e^L$. Furthermore, there exist $w^L > w^H$ such that the principal chooses wages w^θ, for $\theta \in \{H, L\}$, in periods $1, 2, ...$ if she has only observed efforts e^θ and has chosen w^* in period 1 and w^θ thereafter, and she chooses zero wages after any other history. Agent $\theta \in \{H, L\}$ sticks to the effort e^θ as long as the initial wage was w^*, and only w^θ and e^θ have been chosen thereafter. After any history such that the agent of type L has always chosen e^H

and the principal has chosen w^* in period 0 and w^H thereafter, this agent continues to choose e^H. After all other histories, the agents choose minimal efforts. In Appendix A.2.2, we provide a formal definition of an STE and prove the following result.

Proposition 2.3. *In the RPA game with incomplete information, an STE sustaining e^H and e^L, exists for suitable wages w^H and w^L given δ if*

$$\delta S(e^H) \geq c\left(e^H, H\right), \tag{2.5}$$

$$\delta S(e^L) \geq c\left(e^L, L\right) - c\left(e^H, L\right) + c\left(e^H, H\right). \tag{2.6}$$

Proof of Proposition 2.3. See Appendix A.2.2. □

It is simple to show that conditions (2.5) and (2.6) hold for many different choices of e^H and e^L if δ is in a sufficiently small neighbourhood of 1. For instance, for the cost functions used in the experiment, (2.5) and (2.6) both hold with strict inequality for $e^H = 1$ and $e^L = 10$ if δ is sufficiently close to 1. Thus, even though observed behaviour in treatment PI shows no evidence of separation, it is a clear theoretical possibility. Thus, we should think of our empirical result as a statement that, among different possible equilibria, players select those involving pooling behaviour.

2.6 Conclusion

In this chapter, we have shown experimentally that asymmetric information and suspicion about the other player taking an unfair advantage do not necessarily harm the economic performance of relational contracts. Within our main treatments PI and PC, we observe a negative effect of asymmetric information on efficiency. This result mainly follows because dishonest low cost agents imitate the behaviour of true high cost agents. High costs, however, generate inferior levels of efficiency compared to low costs. As many low cost agents choose to be dishonest about their cost type, efficiency in relationships with low cost agents suffers. By contrast, our evidence shows no loss in efficiency as a consequence of asymmetric information within our additional treatments AI and AC. This result is driven by the fact that dishonest low cost agents behave as if they had not lied and provide no different levels of efficiency than low cost agents under complete information. This evidence suggests that asymmetric information may not even harm relational contracts and the provision of efficiency.

Chapter 3

Relational Contracts and the Order of Moves

This chapter shows experimentally that the order of moves within a stage game can be utilized to enhance the economic performance of relational contracts. Repeated principal-agent relationships characterized by stage games in which the agent moves first and the principal thereafter generate higher levels of efficiency compared to relationships in which the principal acts as the first- and the agent as the second-mover in every stage game. Behaviour observed in treatments in which the order of moves alternates after every round, such that the first-mover in the current round becomes the second-mover in the subsequent round, further corroborate that efficiency can be enhanced if the agent rather than the principal moves first within a stage game. The evidence can be organized by a notion of narrow bracketing: despite the ongoing nature of a relationship, second-movers claim more than an equal share of the surplus in a stage game and, as a consequence, principals moving first have to incur much larger costs to induce a given level of efficiency than agents who act in the position of the first-mover.

3.1 Introduction

Many ongoing economic relationships are governed by implicit and informal agreements (see e.g. MacLeod (2007) and Malcomson (2012b) for surveys). Such relational "contracts" are built and maintained by discretionary rewards and punishments executed over the series of transactions which characterizes the repeated relationship. A single transaction often takes place sequentially, particularly in the context of a principal-agent relationship where a stage game within the repeated interaction can consist of the principal paying a wage upfront and the agent exerting effort thereafter. Alternatively, the order of moves can be reversed such that the agent first provides effort and the principal pays a wage thereafter in order to conduct a transaction. This chapter explores how the order of moves in a stage game can be utilized to enhance the provision of efficiency through relational contracts.

The literature views relational contracts, which facilitate cooperation to transact efficiently, primarily as incentivized by the shadow of the future: principal and agent abide by the relational contract if they believe that deviation from the agreement to one's own short-term benefit will be punished in the course of future interaction and therefore not pay out in the long run. If behaviour in relational contracting environments is only driven by the shadow of the future, then the provision of efficiency should not vary with the order of moves within a stage game because, if the relationship is ongoing, then the possibilities to punish deviations from the agreement do not depend on whether a party acts as the first- or second-mover within a stage game. However, this hypothesis does not take into account that individuals may have a narrow rather than broad perspective on their decisions (see e.g. Read et al. (1999) for a survey). In an ongoing relationship, the parties may "narrowly bracket" each stage game to the extent that they, at least partly, consider each transaction in isolation from the subsequent interaction. The outcome of a given stage game may therefore be particularly salient and second-movers may shape it to their advantage by claiming more than an equal share of the surplus within a transaction. Consequently, a principal moving first may be required to invest much larger amounts into a transaction with a given efficiency level than an agent moving first: the wages necessary to contribute to a given surplus may substantially exceed the corresponding costs of effort. A principal who acts as the first-mover may therefore be less willing to induce a transaction with a given efficiency level than an agent in the position of the first-mover.

To investigate these predictions, we study repeated principal-agent relationships in an experiment which builds on the design of Brown et al. (2004, 2012). A relationship is characterized by a sequence of 15 rounds each in which the the principal and her agent conduct a transaction. In every round, an agent exerts costly effort and a principal provides a wage payment. A principal's payoff per transaction increases in effort and an agent's profit per round grows in wage. Each transaction takes place sequentially where the order of moves constitutes the main treatment variable. Under the order "P", the principal provides a wage first and the agent exerts effort thereafter. By contrast, the order "A" dictates the agent to choose his effort first and the principal to pay a wage in the position of the second-mover.

The main result of this chapter shows that the order of moves can be exploited to enhance the economic performance of relational contracts. Relationships in which every transaction is conducted in the order A, where the agent moves first, generate more efficiency than relationships in which the parties transact in the order P, where the principal moves first. The observed behaviour is consistent with the prediction of narrow bracketing as second-movers earn higher profits per transaction than first-movers. This finding is also reflected in the relationship between wages and efforts. In particular, wages for a given effort level are substantially higher if transactions are conducted under the order P than under the order A. As the wages paid by the principals moving first strongly exceed the costs of the corresponding efforts, they have to invest much more than agents in the position of the first-mover to conduct a transaction with a given surplus. If first-movers hesitate to invest large amounts, then this difference may explain why the order P generates lower levels of efficiency than the order A.

To explore the robustness of these patterns, we additionally conducted treatments in which the order of moves alternates between the order P and the order A after every transaction. For instance, if the order A applies to one round, then the subsequent round takes place in the order P, the round after in order A, and so on. In the context of these treatments, the effect of the order of moves on efficiency is strikingly confirmed: transactions conducted under the order A generate higher efficiency than stage games under the order P. Moreover, second-movers in the treatments in which the order alternates reap larger profits from a transaction than first-movers, again consistent with the prediction of narrow bracketing. Hence, the evidence suggests that the order of moves need not persist over the relationship for narrow bracketing and the consequences of the order of moves on efficiency to emerge.

This chapter is related to the experimental literature on relational contracting, particularly to Brown et al. (2004, 2012).[1] In the context of these authors' experiments, the matching of principals and agents as well the duration of their relationships is determined within a market. The authors show that long-term bilateral relationships governed by relational contracts emerge endogenously in markets where explicit contracts, enforced by third parties, are infeasible.[2] The present study builds on this finding and adapts the design proposed by Brown et al. (2004, 2012) to the extent that principals and agents do not participate in a market but instead are randomly assigned into bilateral relationships which last for a fixed number of rounds.

Individuals taking a narrow versus a broad perspective on their decisions have been addressed in the literature on decision-making under risk, in particular, and closest to the

[1]Further experimental research on relational contracting is reported in e.g. Wu and Roe (2007), Fehr et al. (2009) and Camerer and Linardi (2010).

[2]Brown et al. (2012) show that multi-period relationships emerge irrespective of whether the number of principals exceeds or falls below the number of agents in the market. However, more long-term relationships were observed under excess supply of compared to excess demand for agents' labour.

present chapter, with regards to myopic loss aversion (e.g. Samuelson (1963), Benartzi and Thaler (1995), Gneezy and Potters (1997), Thaler et al. (1997), Langer and Weber (2001, 2005), Gneezy et al. (2003), Haigh and List (2005), Sutter (2007), Haisley et al. (2008)).[3] These authors show theoretically and experimentally that subjects' decision-making under risk can depend on the timeframe over which they evaluate outcomes. This insight is related to the hypothesis suggested in the present study that individuals narrowly bracket each transaction rather than focus on the outcome of the relationship as a whole. Despite this similarity, the present study differs from this previous literature because bracketing within the present environment relates to the grouping of decisions in the context of a repeated strategic interaction rather than the bracketing of decisions over lotteries.

At a more general level, this chapter can also be linked to the experimental literature which studies cooperation in repeated two-player games.[4] In particular, Bó (2005) provides strong evidence supporting the view that the shadow of the future affects the provision of efficiency. However, this literature has primarily focussed on behaviour in the repeated prisoner's dilemma.[5] By contrast, the present research concerns relationships with sequential move stage games and players who contribute asymmetrically to efficiency, analogous to many real world relationships. By investigating this setup rather than the repeated prisoner's dilemma and by exogenously varying the sequence of moves, the present study can reveal to what extent the order of moves constitutes a determinant of efficiency.

The remainder of this chapter is structured as follows. Section 3.2 describes the experimental design in detail. The behavioural predictions are addressed in Section 3.3 and the experimental results are presented in Section 3.4. Section 3.5 concludes.

3.2 Experimental Design

3.2.1 Treatments

The present experimental design builds on the setup of Brown et al. (2004), as noted in the introduction. It centres on the bilateral relationship between a principal and her agent, which consists of 15 rounds of interaction. Within each round, a transaction is realized sequentially. The order of moves within a round constitutes this study's main treatment variable. The order of moves P prescribes that a principal first pays a wage and indicates a desired effort level, where the principal's wage payment is binding. The agent observes his principal's choice and then selects an effort, where the desired effort indicated by the

[3] The studies on narrow bracketing and narrow decision framing in the context of decision making under risk including e.g. Tversky and Kahneman (1981), Kahneman and Lovallo (1993), Barberis et al. (2006), Rabin and Weizsäcker (2009) are more broadly related. See Read et al. (1999) for a survey.

[4] See e.g. Roth and Murnighan (1978), Murnighan and Roth (1983), Feinberg and Husted (1993), Palfrey and Rosenthal (1994), Aoyagi and Fréchette (2009), Duffy and Ochs (2009), Bó and Fréchette (2011), Blonski et al. (2011), Fudenberg et al. (2012).

[5] Engle-Warnick and Slonim (2004, 2006a,b) constitute a notable exception. These authors study repeated trust games and focus on the difference between behaviour in finitely and indefinitely repeated games, on the inference of strategies from observed behaviour and on the consequences of the length of the current interaction on trust and trustworthiness in subsequent relationships.

principal is non-binding for the agent's choice. By contrast, in order of moves A, the agent first exerts effort and proposes a desired wage payment. After the principal learns her agent's choice, she selects a wage payment. The agent's desired wage is not binding for the principal in the order A. Taken together, the agent acts as the first- and the principal as the second-mover under the order A, whereas the principal moves first and the agent thereafter under the order P.

The persistence of the order of moves over the course of a relationship constitutes an additional treatment variable within the present design. In the fixed order conditions F, the sequence of moves is the same at each of the 15 rounds of a relationship. In the alternating order conditions X, by contrast, the sequence of moves alternates after every round between the order P and the order A. Taken together, the experiment features a two by two design consisting of the treatments PF, AF, PX and AX, as illustrated in Table 3.1. In particular, every transaction in treatment PF takes place in the order P. By contrast, the order A applies to each round under condition AF. In the treatment PX, order P applies to the first transaction, the order A to the second, the order P to the third, the order A to the fourth round and so on. A relationship conducted under condition AX begins in the order A and continues in the order P in the second round, the order A in the third round, the order P in the fourth round and so on.

Table 3.1: Treatments and the Order of Moves

Treatment	Round 1	Round 2	Round 3	Round 4	...
PF	P	P	P	P	...
AF	A	A	A	A	...
PX	P	A	P	A	...
AX	A	P	A	P	...

3.2.2 Parameters

Within every round, a principal chooses a wage w from $[0, 100]$ and an agent selects an effort e from $\{1, \ldots, 10\}$. The desired wage and desired effort are selected from the same sets as wages and efforts, respectively. For given choices of w and e, the payoff for the principal $\Pi_P(w, e)$ and the payoff for the agent $\Pi_A(w, e)$ within a round is computed as follows:

$$\Pi_P(w, e) = 10 \cdot e - w \text{ and } \Pi_A(w, e) = w - c(e),$$

where $c(e)$ denotes the cost of effort e as summarized in Table 3.2 below. The costs are strictly increasing and the marginal costs are weakly increasing in effort. However, since the marginal costs at every feasible effort level are lower than the marginal benefit, the highest possible effort level induces efficiency.

Table 3.2: Cost of Effort

e	1	2	3	4	5	6	7	8	9	10
$c(e)$	0	3	6	10	15	20	25	30	36	42

3.2.3 Procedures

At the beginning of the experiment, the subjects are randomly assigned to either the role of a principal or the role of an agent and the assignment remains fixed throughout the experiment. Principals and agents are randomly matched and the matching persists for the fifteen rounds of a relationship. The interaction between a principal and an agent is framed as a buyer-seller relationship.[6] At the end of each round, every subject receives a summary of his action(s), the behaviour of the subject he is matched with and his own payoff in the current round. This summary of results contains the same elements across all treatment conditions. During the experiment, payoffs are noted in terms of the experimental currency “Punkte” (points). The subjects are provided with a show up fee of 100 points.

Every session of a given treatment is executed under the same protocol. Yet, the procedures for fixed and alternating order treatments differ in the number of matches between principals and agents. In every fixed order treatment, a subject exclusively interacts with one other subject and every payoff earned during the match is converted into real money by the end of the experiment (10 Punkte = 1 CHF). In the context of every alternating order treatment, each subject participates in a total of four matches. After the fifteenth round of a given match has ended, principals and agents are rematched using a perfect stranger protocol.[7] One of the four matches is randomly selected and every payoff earned during this match is paid out by the end of the experiment (10 Punkte = 1 CHF). Unless noted otherwise, the data presented in the following section correspond to behaviour elicited within each subject’s first match.[8]

The experiment was computerized using the software z-tree (Fischbacher, 2007). The software hroot (Bock et al., 2012) was employed for organizing and recruitment. The subject pool primarily consists of students at the University of Zurich and the Swiss Federal Institute of Technology in Zurich. A total of 188 subjects participated in the experiment which took place during October 2013, April and May 2014. The number of sessions per

[6]Sample instructions are provided in Appendix B, Section B.2.

[7]In particular, subjects in a given session are randomly assigned into matching groups, each consisting of four principals and four agents. Over the course of the session, every principal interacts with every agent in her matching group for one fifteen period relationship.

[8]An additional procedural difference between the alternating and fixed order treatments concerns the presentation of the effort costs. In every alternating order condition, costs are presented as listed in Table 3.2. Under the fixed order protocol, every agent is randomly assigned to either a high or low cost structure (see the sample instructions in Appendix B, Section B.2.1, for the exact specification). High costs are equivalent to the costs listed in Table 3.2. Low costs are lower than high costs and featured lower marginal costs for every given effort level. The present chapter is confined to data elicited under high costs. Hence, every agent in the alternating and the fixed order conditions faces the same costs and the differences between these conditions relate to the labelling of the costs and the possibility that an agent could be assigned to a different cost structure within a fixed but not within an alternating order treatment.

treatment and the number of subjects participating in a treatment are summarized in Table 3.3. A session lasted 105 minutes on average. Subjects' average earnings in the context of the experiment amount to 40.6CHF (43.4$).

Table 3.3: Number of Sessions and Subjects per Session

Treatment	Sessions	Subjects per Treatment
PF	2	28
AF	2	32
PX	2	64
AX	2	64

3.3 Predictions

Principals and agents within our experimental environment cannot enter into binding agreements. Hence, the parties face short term incentives to behave selfishly rather than cooperatively, for instance by provision of minimal efforts and wages rather than efficient effort and fair wage payments. As an extensive literature has argued, however, relational contracts, informal agreements which emerge in the context of an ongoing relationship, can provide incentives to cooperate, even in the absence of formal contracting opportunities (see e.g. MacLeod (2007) and Malcomson (2012b) for surveys). In particular, relational contracts can be enforced by the shadow of the future: the parties abide by the implicit agreement because they expect the relationship to turn sour otherwise. For instance, a principal and her agent may implicitly agree to cooperate if the short run benefits of shirking are outweighed by the costs of being punished in the future for not having honoured the agreement.

If a relationship is ongoing, i.e. within the first fourteen rounds of a relationship in the experimental environment, then the possibilities to punish a breach on a relational contract do not depend on whether the infringing party moves first or second within a stage game. For instance, a downward deviation from the implicitly agreed effort level in the current round can be punished by a subsequent reduction in the wage payment - either in the current round if the deviating agent moves first, or in the subsequent round if the shirking agent acts as the second-mover in the current round. As a consequence, an agent moving first should provide similar levels of effort than an agent moving second because selfish behaviour can be punished in either case. Analogously, wage payments should not depend on whether a principal moves first or second in an ongoing relationship. Taken together, if the shadow of the future is the sole driver of behaviour, then the order of moves within a stage game should not affect the provision of efficiency.

Hypothesis 3.1. *Efficiency does not depend on the order of moves within a stage game.*

Analogous to many real world repeated interactions, our experimental design frames a relationship as consisting of a sequence of transactions. Despite the ongoing nature of the relationship, principals and agents may *narrowly bracket* each transaction: they focus on the current stage game and, to some extent, neglect subsequent interaction in the relationship. The subjects may especially concentrate on the outcome within a transaction, to which the first-mover contributes by investing in advance whereas the second-mover's role lies in the distribution of payoffs. Such narrow bracketing of a stage game could have consequences for the provision of efficiency.

If first-movers consider themselves as investing in advance into a transaction, then they may give particular attention to the possibility that second-movers may not honour their investments. Observe that principals have to contribute larger amounts than agents in order to conduct a transaction with a particular surplus and sharing rule.[9] For instance, if the parties want to implement an efficient surplus and a sharing rule which splits the surplus equally between the parties, then the agent has to incur the costs of efficient effort, which equal 42 points, and the principal has to pay the fair wage for efficient effort, which amounts to 71 points. This may imply that principals moving first under the order P may be less willing to induce a transaction with a given efficiency level compared to agents moving first under the order A. This line of argument suggests that efficiency provided under the order P could be lower than under the order A.

Second-movers, by definition, decide on the distribution of payoffs within a given round. If narrow bracketing implies a particular focus on the outcome in a given round, then second-movers may influence the outcome to their own payoff advantage. This suggests that second-movers earn higher profits from a transaction than first-movers. Such effect could contribute to the consequences of the order of moves on efficiency as proposed above. For instance, if agents who act as second-movers claim more than an equal share of the surplus, then their principals are forced to pay more than fair wages as first-movers to achieve a given surplus. Higher than fair wages can substantially exceed the costs of effort necessary to implement a given level of efficiency. As a consequence, principals moving first under the order P may be less willing to invest into a transaction with a given surplus than agents moving first under the order A. Hence, second-movers who narrowly bracket and distribute the payoffs within a round to their own advantage could have a hand in the effect of the order of moves on efficiency.

Hypothesis 3.2. *The order of moves affects the provision of efficiency: the order P leads to less efficiency than the order A.*

The following section presents the empirical results obtained from the experimental design. As the above stated predictions hinge on the assumption that a relationship is ongoing, the subsequent analysis will focus on behaviour elicited within the first fourteen rounds, in which a relationship continues after every transaction. However, the empirical

[9] More formally, fix a surplus generated by effort level e and let α denote the agent's and $(1-\alpha)$ the principal's share of the surplus, where $\alpha \in (0,1)$. To implement such a transaction, the agent has to invest $c(e)$ and the principal has to pay $\alpha 10e + (1-\alpha)c(e)$. It is straightforward to see that the latter is strictly lower than the former for any given $e \in \{1,\dots,10\}$ and $\alpha \in (0,1)$. Thus, for a fixed surplus and sharing rule, the costs of investment for the principal always exceed the agent's.

effects reported below are documented in almost all rounds of the observed relationships. Hence, alternative specifications of the investigated time horizon would lead to very similar results.

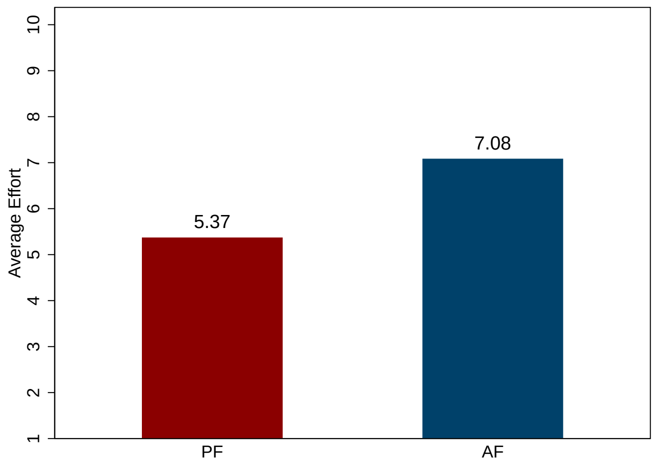
3.4 Results

To explore the consequences of the sequence of moves on efficiency, behaviour in the conditions PF and AF will be compared. These two treatments differ in the order of moves: in PF, the principal moves first and the agent thereafter in each stage game whereas the agent moves first and the principal thereafter at every transaction taking place in treatment AF. Hypothesis 3.1 stated above predicts no difference in efficiency between these treatments. By contrast, Hypothesis 3.2 suggests that effort provided in PF is lower than in AF.

Result 3.1. *The order of moves affects efficiency: effort in treatment PF is lower than effort in treatment AF.*

The empirical data support Hypothesis 3.2. As Figure 3.1 shows, average effort provided in treatment PF equals 5.37 and is significantly different from average effort provided in treatment AF, which equals 7.08 (p-value 0.058, Wilcoxon rank-sum test[10]). The difference in efforts between the two treatments already exists in the first rounds and continues over the course of the relationships.[11]

Figure 3.1: Effort in the Fixed Order Treatments



Note: Average efforts are computed over the first fourteen rounds of a given relationship and over all relationships within a given treatment.

Hypothesis 3.2 has been motivated by the assertion of narrow bracketing. To explore if our data is consistent with the hypothesized characteristics of bracketing behaviour, we address subjects' profits from each transaction. If the subjects narrowly bracket to the

[10]This test is performed on agents' average efforts in the fixed order treatments PF and AF where the average for each agent is computed over the first fourteen rounds of his relationship.
[11]See Figure B.1 in Appendix B.1.

extent that they focus on the outcome of the current transaction and, at least partially, disregard the subsequent interaction, then second-movers may exploit their role as deciding on the distribution of payoffs. In particular, second-movers may claim higher profits for themselves than allocate to their first-movers. Yet, second-movers earning higher profits compared to their first-movers may be considered uncooperative and therefore be punished in future rounds. In anticipation of this effect, second-movers may refrain from exploiting their role to the own advantage. However, this hypothesis is not supported by our data.

Result 3.2. *Second-movers earn higher profits per transaction than first-movers in the fixed order treatments PF and AF, consistent with the hypothesis of narrow bracketing.*

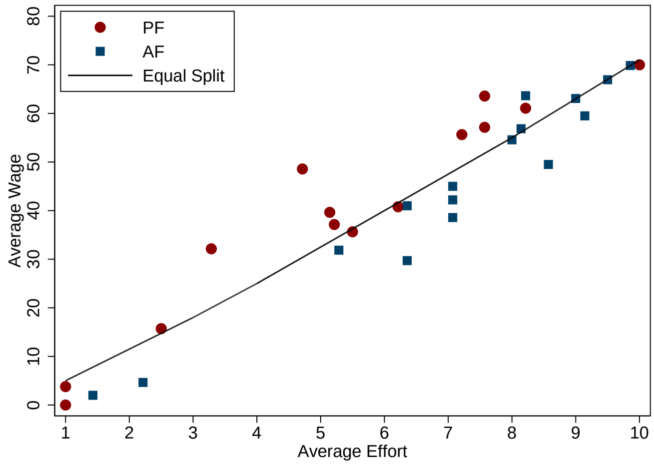
Considering treatment PF, average profits per round earned by the principals equal 13.6 points, which significantly differ from agents' average profits per transaction, amounting to 21.5 points (p-value 0.06, Wilcoxon signed-rank test[12]). In treatment AF, by contrast, principals' average profits per round, 25.9 points, are significantly different from agents' average profits per transaction, 18.0 points (p-value 0.006, Wilcoxon signed-rank test). Beginning with the first stage game, the difference between first- and second-mover profits is persistent over the rounds.[13]

The relationship between the wages and efforts exerted within the relationships provides further insight concerning the observed asymmetry in payoffs. A regression analysis indicates that wages for a given effort level in treatment PF are on average 8.76 points per transaction higher than in treatment AF.[14] As Figure 3.2 illustrates, almost all outcomes of relationships observed in treatment PF lie weakly above the line indicating the equal split of the surplus. In other words, almost all principals pay higher than fair wages, if they act in the position of the first-mover. On the contrary, almost all outcomes of relationships in the context of treatment AF lie weakly below the equal split line. Thus, almost all principals pay less than fair wages if they act as second-movers.

The evidence shows that principals in condition PF invest substantially higher amounts as first-movers to induce a given efficiency level compared to agents moving first in condition AF: the wages for a given effort level paid by the principals in PF strongly exceed the corresponding effort costs which agents incur as first-movers in AF. This difference in first-mover investments can be linked to the effect of the order of moves on efficiency: if first-movers hesitate to invest large amounts into a transaction because their second-movers cannot commit to honouring it, then the observed difference in first-mover investments between principals and agents may explain why relationships in condition PF settle on outcomes with lower efficiency than relationships in condition AF.

[12]In the following, Wilcoxon signed-rank tests on profits are performed on subjects' average profits, where the average profit for each individual is taken over the first fourteen rounds of his relationship.
[13]See Figure B.2 in Appendix B.1.
[14]See Appendix B.1, Table B.1, Column (1).

Figure 3.2: Wage-Effort Relationship in the Fixed Order Treatments



Notes: One data point marked with a dot (square) in the above plot represents the vector of average wage and average effort generated by a principal-agent pair in the treatment PF(AF), where the average is computed over the first fourteen rounds of the relationship. "Equal Split" indicates the wage which splits the surplus equally between the parties for every given effort, where linear interpolation is applied.

To investigate the robustness of the above reported findings, behaviour observed in the conditions PX and AX will be addressed below. The first round of a relationship in PX takes place in the order P, whereas the first round in AX is conducted in the order A. Thereafter, the order of moves alternates in the treatments PX and AX to the extent that a stage game in the order P follows a transaction in the order A and vice versa. The analysis of behaviour observed in the alternating order conditions therefore indicates to what extent the persistence of the order of moves in the treatments PF and AF is necessary for the consequences of the order of moves on efficiency to emerge. In case of the alternating order treatments PX and AX, Hypothesis 3.2 predicts that transactions conducted under the order P are characterized by lower efficiency than transactions taking place under the order A. According to Hypothesis 3.1, no such difference should be observed.

Result 3.3. *(i) The order of moves affects efficiency in the alternating order treatments PX and AX: effort under the order P is lower than effort in the context of the order A.*
(ii) Second-movers earn higher profits per transaction than first-movers in the alternating order treatments, providing further support for the hypothesis of narrow bracketing.

Considering relationships in the alternating order treatments, average effort exerted in stage games under the order P equals 6.27 and is significantly different from average effort in transactions executed under the order A, 7.37 (p-value 0.000, Wilcoxon signed-rank test[15]). Differences in efforts depending on the order of moves exist in almost all rounds of

[15] This test is performed on average efforts exerted by agents participating in the treatments PX and AX. For each of these agents, one average is computed over all transactions conducted under the order P and a second average is calculated over all stage games in the order A where, in both cases, the first fourteen stage games are taken into consideration.

the relationships.[16] Thus, these findings confirm the consequences of the order of moves on efficiency. Moreover, the evidence suggests that for such effects to arise, the order of moves need not be the same in every transaction of a relationship. This illustrates the robustness of the pattern observed in Result 3.1.

Figure 3.3: Effort in the Alternating Order Treatments

Note: Average effort in column "Transaction in Order P(A)" is computed over all relationships in the alternating order conditions PX and AX, where all transactions executed in the order P(A) within the first fourteen rounds are taken into consideration.

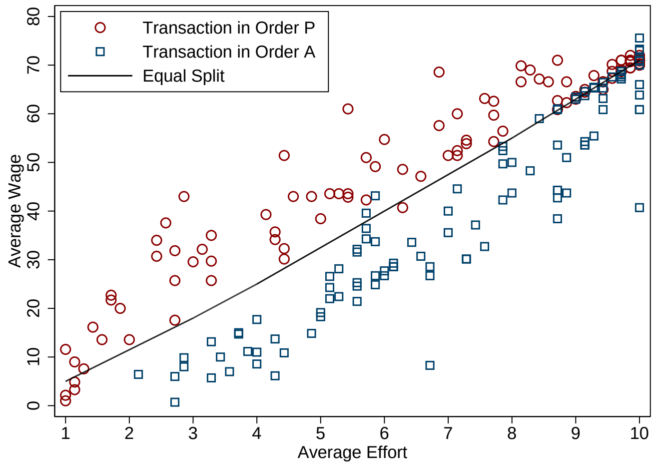
We further investigate whether behaviour in the alternating order treatments shows signs of narrow bracketing. In particular, we ask if asymmetries in profits between the first- and second-mover go along with the observed effect on efficiency. Considering transactions in the alternating order treatments conducted in the order P, principals earn 15.6 points on average per round, which is significantly different from agents' per transaction earnings, 26.3 points (p-value 0.00, Wilcoxon signed-rank test). With regards to stage games in the order A, principals' average earnings per transaction, 31.7 points, are significantly different from agents' per transaction profits, 14.5 points (p-value 0.00, Wilcoxon signed-rank test). The difference in profits between first- and second-movers persists over the course of the relationships.[17] This evidence confirms the pattern observed in Result 3.2.

The relationship between wages and efforts sheds further light on the asymmetry in parties' profits in the alternating order treatments. A regression analysis shows that, in these conditions, wages for a given effort level are on average 16.25 points higher if a transaction takes place in the order P compared to the order A.[18] As Figure 3.4 illustrates, almost all outcomes of transactions conducted under the order P are located above the line which indicates the equal split of the surplus between the parties. In other words, almost all principals pay higher than fair wages under the order P. This implies that principals moving first in the order P invest substantially higher amounts to induce a transaction with

[16]See Figure B.3 in Appendix B.1.
[17]See Figure B.4 in Appendix B.1.
[18]See Appendix B.1, Table B.1, Column (2).

a given efficiency level than agents moving first under the order A. The evidence therefore provides further support for the intuition that narrow bracketing and the observed effect of the order of moves on efficiency are related.

Figure 3.4: Wage-Effort Relationship in the Alternating Order Treatments



Notes: One data point marked with a circle (square) represents the vector of the average wage and average effort provided by one principal-agent relationship observed in either treatment PX or treatment AX, where the average is computed over all stage games conducted in the order P(A) within the first fourteen rounds of the relationship. "Equal Split" indicates the wage which splits the surplus equally between the parties for every given effort where linear interpolation is applied.

3.5 Conclusion

This chapter illustrates how the order of moves within a stage game can be used to enhance the provision of efficiency in repeated principal-agent relationships governed by relational contracts. The previous literature primarily views relational contracts as incentivized by the shadow of the future: the parties stick to an implicit agreement to cooperate if they believe that deviating to one's own short term benefit will be punished in the course of the future interaction. The intuition that repeated interaction can provide incentives for cooperation is consistent with our data as average efficiency across all treatments is higher than minimal efficiency. However, the rationale that the shadow of the future incentivizes efficient behaviour does not predict the order of moves within a stage game to matter for the provision of efficiency. This view therefore cannot explain the robust finding in our experiment that transactions in which the agent moves first generate higher levels of efficiency than stage games in which the principal acts as the first-mover. We hypothesize that a notion of narrow bracketing can organize the observed pattern. In particular, the data show that second-movers affect the outcome of the current stage game to their own advantage. This suggests that they narrowly focus on the current round rather than on the relationship as a whole. As a consequence, principals moving first have to invest higher amounts than agents moving first in order to conduct a transaction with a given level of efficiency. This may explain why stage games in which the principal rather than the agent moves first lead to less efficiency.

Several distinct mechanisms could have caused the bracketing behaviour. As Read et al. (1999) note, individuals may narrowly bracket because of cognitive capacity limitations. Within the present environment, second-movers may not fully comprehend the negative consequences of claiming more than an equal share of the surplus for the provision of efficiency and for their potential profits from the overall relationship. Moreover, individuals may narrowly bracket each transaction to achieve a certain goal, which Read et al. (1999) refer to as motivated bracketing. In the context of our experiment, individuals may want to earn higher profits than their relational partner. Such kind of preference is consistent with the observation that second-movers secure higher profits for themselves than allocate to their first-movers. Investigating the relative impact of these mechanisms on the observed behaviour constitutes only one out of the many interesting roads for future research which the reported effects have unveiled.

Bibliography

Alger, I. and Renault, R. (2006). Screening ethics when honest agents care about fairness. *International Economic Review*, 47(1):59–85.

Andreoni, J., Brown, P., and Vesterlund, L. (2002). What makes an allocation fair? Some experimental evidence. *Games and Economic Behavior*, 40(1):1–24.

Antler, Y. (2014). Two sided matching with endogenous preferences. *American Economic Journal: Microeconomics*, forthcoming.

Aoyagi, M. and Fréchette, G. (2009). Collusion as public monitoring becomes noisy: experimental evidence. *Journal of Economic Theory*, 144(3):1135–1165.

Arrow, K. (1979). The property rights doctrine and demand revelation under incomplete information. In Boskin, M. J., editor, *Economics and Human Welfare*. Academic Press.

Bó, P. (2005). Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *American Economic Review*, 95(5):1591–1604.

Bó, P. and Fréchette, G. (2011). The evolution of cooperation in infinitely repeated games: experimental evidence. *American Economic Review*, 101(1):411–429.

Barberis, N., Huang, M., and Thaler, R. (2006). Individual preferences, monetary gambles, and stock market participation: a case of narrow framing. *American Economic Review*, 96(4):1069–1090.

Bartling, B. and Netzer, N. (2014). An externality-robust auction: theory and experimental evidence. Mimeo.

Bassi, M., Pagnozzi, M., and Piccolo, S. (2014). Optimal contracting with altruism and reciprocity. *Research in Economics*, 68(1):27–38.

Battigalli, P. and Dufwenberg, M. (2009). Dynamic psychological games. *Journal of Economic Theory*, 144(1):1–35.

Benartzi, S. and Thaler, R. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics*, 110(1):73–92.

Bergemann, D. and Morris, S. (2005). Robust mechanism design. *Econometrica*, 73(6):1771–1813.

Bierbrauer, F. and Netzer, N. (2011). Mechanism design and intentions. Mimeo.

Bierbrauer, F. and Netzer, N. (2014). Mechanism design and intentions. University of Zurich, Department of Economics, Working Paper No. 66.

Bierbrauer, F., Ockenfels, A., Pollak, A., and Rückert, D. (2014). Robust mechanism design and social preferences. Mimeo.

Blonski, M., Ockenfels, P., and Spagnolo, G. (2011). Equilibrium selection in the repeated prisoner's dilemma: axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics*, 3(3):164–192.

Blount, S. (1995). When social outcomes aren't fair: the effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes*, 63(2):131–144.

Bock, O., Nicklisch, A., and Baetge, I. (2012). hroot: Hamburg registration and organization online tool. H-Lab Working Paper No.1.

Bodoh-Creed, A. (2012). Ambiguous beliefs and mechanism design. *Games and Economic Behavior*, 75(2):518–537.

Bose, S., Ozdenoren, E., and Pape, A. (2006). Optimal auctions with ambiguity. *Theoretical Economics*, 1(4):411–438.

Bowles, S. and Polanía-Reyes, S. (2012). Economic incentives and social preferences: substitutes or complements? *Journal of Economic Literature*, 50(2):368–425.

Brown, M., Falk, A., and Fehr, E. (2004). Relational contracts and the nature of market interactions. *Econometrica*, 72(3):747–780.

Brown, M., Falk, A., and Fehr, E. (2012). Competition and relational contracts: the role of unemployment as a disciplinary device. *Journal of the European Economic Association*, 10(4):887–907.

Cabrales, A. and Serrano, R. (2011). Implementation in adaptive better-response dynamics: towards a general theory of bounded rationality in mechanisms. *Games and Economic Behavior*, 73(2):360–374.

Camerer, C. and Linardi, S. (2010). Can relational contracts survive stochastic interruptions? Experimental evidence. University of Pittsburgh, Department of Economics, Working Paper No. 483.

Charness, A. and Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117(3):817–869.

Cox, J., Friedman, D., and Gjerstad, S. (2007). A tractable model of reciprocity and fairness. *Games and Economic Behavior*, 59(1):17–45.

Crawford, V., Neeman, Z., Kugler, T., and Pauzner, A. (2009). Behaviorally optimal auction design: examples and observations. *Journal of the European Economic Association*, 7(2-3):377–387.

Crémer, J. and McLean, R. (1985). Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53(2):345–361.

d'Aspremont, C. and Gerard-Varet, L.-A. (1979). Incentives and incomplete information. *Journal of Public Economics*, 11(1):25–45.

de Clippel, G. (2014). Behavioral implementation. *American Economic Review*, forthcoming.

De Marco, G. and Immordino, G. (2012). Reciprocity in the principal multiple agent model. University of Naples, Department of Economics, CSEF Working Paper No. 314.

De Marco, G. and Immordino, G. (2013). Partnership, reciprocity and team design. *Research in Economics*, 67(1):39–58.

Desiraju, R. and Sappington, D. (2007). Equity and adverse selection. *Journal of Economics and Mangement Strategy*, 16(2):285–318.

Duffy, J. and Ochs, J. (2009). Cooperative behavior and the frequency of social interaction. *Games and Economic Behavior*, 66(2):785 – 812.

Dufwenberg, M., Gächter, S., and Hennig-Schmidt, H. (2011). The framing of games and the psychology of play. *Games and Economic Behavior*, 73(2):459–478.

Dufwenberg, M. and Kirchsteiger, G. (2000). Reciprocity and wage undercutting. *European Economic Review*, 44(4-6):1069–1078.

Dufwenberg, M. and Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2):268–298.

Dufwenberg, M., Smith, A., and Van Essen, M. (2013). Hold-up: With a vengeance. *Economic Inquiry*, 51(1):896–908.

Eisenhuth, R. (2012). Reference dependent mechanism design. Mimeo.

Eisenhuth, R. and Ewers, M. (2012). Auctions with loss averse bidders. Mimeo.

Eliaz, K. (2002). Fault tolerant implementation. *Review of Economic Studies*, 69(3):589–610.

Engle-Warnick, J. and Slonim, R. (2006a). Inferring repeated-game strategies from actions: evidence from trust game experiments. *Economic Theory*, 28(3):603–632.

Engle-Warnick, J. and Slonim, R. (2006b). Learning to trust in indefinitely repeated games. *Games and Economic Behavior*, 54(1):95–114.

Engle-Warnick, J. and Slonim, R. L. (2004). The evolution of strategies in a repeated trust game. *Journal of Economic Behavior and Organization*, 55(4):553–573.

Englmaier, F. and Segal, C. (2012). Unions, communication, and cooperation in organizations. Mimeo.

Falk, A., Fehr, E., and Fischbacher, U. (2003). On the nature of fair behavior. *Economic Inquiry*, 41(1):20–26.

Falk, A., Fehr, E., and Fischbacher, U. (2008). Testing theories of fairness - intentions matter. *Games and Economic Behavior*, 62(1):287–303.

Falk, A. and Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2):293–315.

Fehr, E., Brown, M., and Zehnder, C. (2009). On reputation: a microfoundation of contract enforcement and price rigidity. *Economic Journal*, 119(536):333–353.

Feinberg, R. and Husted, T. (1993). An experimental test of discount-rate effects on collusive behaviour in duopoly markets. *Journal of Industrial Economics*, 41(2):153–160.

Filiz-Ozbay, E. and Ozbay, E. (2007). Auctions with anticipated regret: theory and experiment. *American Economic Review*, 97(4):1407–1418.

Fischbacher, U. (2007). z-tree - zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Fischbacher, U. and Foellmi-Heusi, F. (2013). Lies in disguise - an experimental study on cheating. *Journal of the European Economic Association*, 11(3):525–547.

Fudenberg, D., Rand, D., and Dreber, A. (2012). Slow to anger and fast to forgive: cooperation in an uncertain world. *American Economic Review*, 102(2):720–49.

Geanakoplos, J., Pearce, D., and Stacchetti, E. (1989). Psychological games and sequential rationality. *Games and Economic Behavior*, 1(1):60–79.

Gibson, R., Tanner, C., and Wagner, A. F. (2013). Preferences for truthfulness: heterogeneity among and within individuals. *American Economic Review*, 103(1):532–548.

Glazer, A. and Rubinstein, A. (1998). Motives and implementation: on the design of mechanisms to elicit opinions. *Journal of Economic Theory*, 79(2):157–173.

Gneezy, U. (2005). Deception: the role of consequences. *American Economic Review*, 95(1):384–394.

Gneezy, U., Kapteyn, A., and Potters, J. (2003). Evaluation periods and asset prices in a market experiment. *Journal of Finance*, 58(2):821–837.

Gneezy, U. and Potters, J. (1997). An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics*, 112(2):631–645.

Hahn, V. (2009). Reciprocity and voting. *Games and Economic Behavior*, 67(2):467–480.

Haigh, M. and List, J. (2005). Do professional traders exhibit myopic loss aversion? An experimental analysis. *Journal of Finance*, 60(1):523–534.

Haisley, E., Mostafa, R., and Loewenstein, G. (2008). Myopic risk-seeking: the impact of narrow decision bracketing on lottery play. *Journal of Risk and Uncertainty*, 37(1):57–75.

Halac, M. (2012). Relational contracts and the value of relationships. *American Economic Review*, 102(2):750–779.

Hoffmann, M. and Kolmar, M. (2013). Intention-based fairness preferences in two-player contests. *Economics Letters*, 120(2):276–279.

Hurkens, S. and Kartik, N. (2009). Would I lie to you? On social preferences and lying aversion. *Experimental Economics*, 12(2):180–192.

Jehiel, P., Meyer-Ter-Vehn, M., Moldovanu, B., and Zame, W. (2006). The limits of ex post implementation. *Econometrica*, 74(3):585–610.

Jehiel, P. and Moldovanu, B. (2006). Allocative and informational externalities in auctions and related mechanisms. In Blundell, R., Newey, W., and Persson, T., editors, *Proceedings of the 9th World Congress of the Econometric Society*. Cambridge University Press.

Kahneman, D. and Lovallo, D. (1993). Timid choices and bold forecasts: a cognitive perspective on risk taking. *Management Science*, 39(1):17–31.

Kucuksenel, S. (2012). Behavioral mechanism design. *Journal of Public Economic Theory*, 14(5):767–789.

Langer, T. and Weber, M. (2001). Prospect theory, mental accounting, and differences in aggregated and segregated evaluation of lottery portfolios. *Management Science*, 47(5):716–733.

Langer, T. and Weber, M. (2005). Myopic prospect theory vs. myopic loss aversion: how general is the phenomenon? *Journal of Economic Behavior and Organization*, 56(1):25–38.

Levin, J. (2003). Relational incentive contracts. *American Economic Review*, 93(3):835–857.

Levine, D. (1998). Modelling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1(3):593–622.

Li, J. and Matouschek, N. (2014). Managing conflicts in relational contracts. *American Economic Review*, 103(6):2328–2351.

Luce, R. D. and Raiffa, H. (1957). *Games and Decisions*. John Wiley and Sons, Inc.

MacLeod, W. B. (2007). Reputation, relationships, and contract enforcement. *Journal of Economic Literature*, 45(3):595–628.

Malcomson, J. (2012a). Relational incentive contracts with persistent private information. University of Oxford, Department of Economics, Economics Series Working Papers No. 634.

Malcomson, J. M. (2012b). Relational incentive contracts. In Gibbons, R. and Roberts, J., editors, *Handbook of Organizational Economics*. Princeton University Press.

Maskin, E. and Riley, J. (1984). Optimal auctions with risk averse buyers. *Econometrica*, 52(2):1473–1518.

Mathevet, L. (2010). Supermodular mechanism design. *Theoretical Economics*, 5(3):403–443.

Mazar, N., Amir, O., and Ariely, D. (2008). The dishonesty of honest people: a theory of self-concept maintenance. *Journal of Marketing Research*, 45(6):633–644.

Mookherjee, D. and Reichelstein, S. (1990). Implemention via augmented revelation mechanisms. *Review of Economic Studies*, 57(3):453–475.

Murnighan, J. K. and Roth, A. E. (1983). Expecting continued play in prisoner's dilemma games: a test of several models. *Journal of Conflict Resolution*, 27(2):279–300.

Myerson, R. and Satterthwaite, M. (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 28(2):265–281.

Netzer, N. and Schmutzler, A. (2013). Explaining gift-exchange – the limits of good intentions. *Journal of the European Economic Association*, forthcoming.

Nishimura, N., Cason, T., Saijo, T., and Ikeda, Y. (2011). Spite and reciprocity in auctions. *Games*, 2(3):365–411.

Palfrey, T. and Rosenthal, H. (1994). Repeated play, cooperation and coordination: an experimental study. *Review of Economic Studies*, 61(3):545–565.

Pruckner, G. and Sausgruber, R. (2013). Honesty on the streets: a field study on newspaper purchasing. *Journal of the European Economic Association*, 11(3):661–679.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83(5):1281–1302.

Rabin, M. and Weizsäcker, G. (2009). Narrow bracketing and dominated choices. *American Economic Review*, 99(4):1508–1543.

Read, D., Loewenstein, G., and Rabin, M. (1999). Choice bracketing. *Journal of Risk and Uncertainty*, 18(1):171–197.

Renou, L. and Schlag, K. (2011). Implementation in minimax regret equilibrium. *Games and Economic Behavior*, 71(2):527–533.

Roth, A. and Murnighan, J. (1978). Equilibrium behavior and repeated play of the prisoner's dilemma. *Journal of Mathematical Psycholog*, 17(2):189–198.

Samuelson, P. A. (1963). Risk and uncertainty: a fallacy of large numbers. *Scientia*, 98:108–113.

Sanchez-Pages, S. and Vorsatz, M. (2007). An experimental study of truth-telling in a sender-receiver game. *Games and Economic Behavior*, 61(1):86–112.

Saran, R. (2011). Menu-dependent preferences and revelation principle. *Journal of Economic Theory*, 146(4):1712–1720.

Sebald, A. (2010). Attribution and reciprocity. *Games and Economic Behavior*, 68(1):339–352.

Segal, U. and Sobel, J. (2007). Tit for tat: foundations of preferences for reciprocity in strategic settings. *Journal of Economic Theory*, 136(1):197–216.

Segal, U. and Sobel, J. (2008). A characterization of intrinsic reciprocity. *International Journal of Game Theory*, 36(3-4):571–585.

Sobel, J. (2005). Interdependent preferences and reciprocity. *Journal of Economic Literature*, 43(2):392–436.

Sutter, M. (2007). Are teams prone to myopic loss aversion? An experimental study on individual versus team investment behavior . *Economics Letters*, 97(2):128–132.

Tang, P. and Sandholm, T. (2012). Optimal auctions for spiteful bidders. *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, pages 1457–1463.

Thaler, R., Tversky, A., Kahneman, D., and Schwartz, A. (1997). The effect of myopia and loss aversion on risk taking: an experimental test. *Quarterly Journal of Economics*, 112(2):647–661.

Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(30):453–458.

von Siemens, F. (2009). Bargaining under incomplete information, fairness, and the hold-up problem. *Journal of Economic Behavior and Organization*, 71(2):486–494.

von Siemens, F. (2011). Heterogeneous social preferences, screening, and employment contracts. *Oxford Economic Papers*, 63:499–522.

von Siemens, F. (2013). Intention-based reciprocity and the hidden costs of control. *Journal of Economic Behavior and Organization*, 92:55–65.

Wu, S. and Roe, B. E. (2007). Discretionary latitude and relational contracting. IZA Discussion Papers, No. 2879.

Yang, H. (2013). Nonstationary relational contracts with adverse selection. *International Economic Review*, 54(2):525–547.

A Appendix to Chapter 2

A.1 Regressions and Figures

Table A.1: Wage-Effort Regressions

	(1)	(2)	(3)	(4)	(5)
Constant	7.84^{***}	6.89^{**}	5.93^{***}	-4.22	-7.09^{***}
	(1.34)	(3.11)	(0.98)	(3.00)	(2.69)
Effort	6.57^{***}	5.92^{***}	6.03^{***}	7.51^{***}	6.69^{***}
	(0.25)	(0.44)	(0.17)	(0.95)	(0.35)
Effort $\cdot\ \mathbf{1}_{[H]}$	-	1.89^{***}	1.25^{***}	-0.37	1.75^{***}
		(0.36)	(0.17)	(0.86)	(0.28)
N	92	28	64	32	32
Adj. R^2	0.88	0.89	0.94	0.82	0.91
Sample	PC and PI	PC	PI	AI	AC

Notes: The dependent variable in all of the above regressions is wage. The analysis is performed on average wages and average efforts (recoded s.t. $e \in \{0, 1, \ldots, 9\}$, where the averages are taken over time for every pair of principal and agent. $\mathbf{1}_{[H]}$ represents a dummy which equals one if costs are reported to be high in the incomplete information conditions (PI and AI) or costs are truly high in the context of the complete information conditions (PC and AC) and zero otherwise. All reported standard errors are robust. ***(**) denotes significance at the 1(5) percent level.

Figure A.1: Effort in Case of Low Costs (PC), and Honesty and Low Costs (PI)

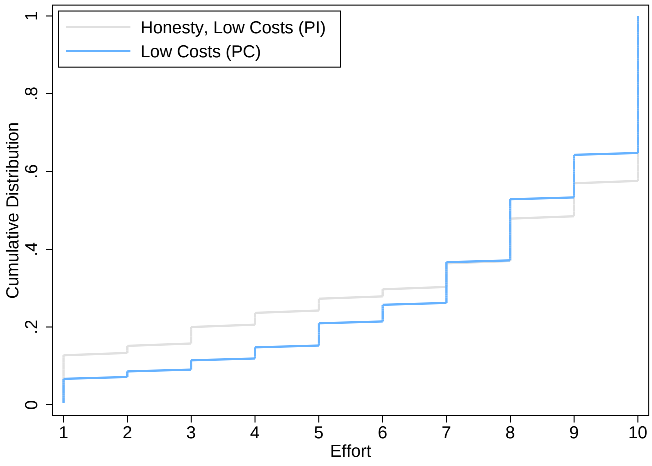


Figure A.2: Effort in Case of Low Costs (PC and PI)

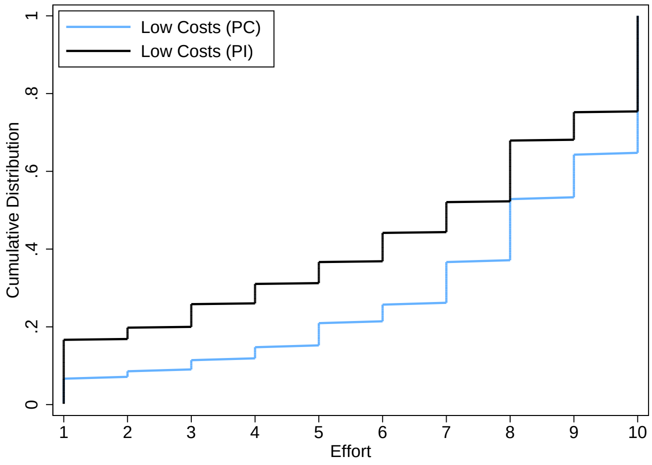
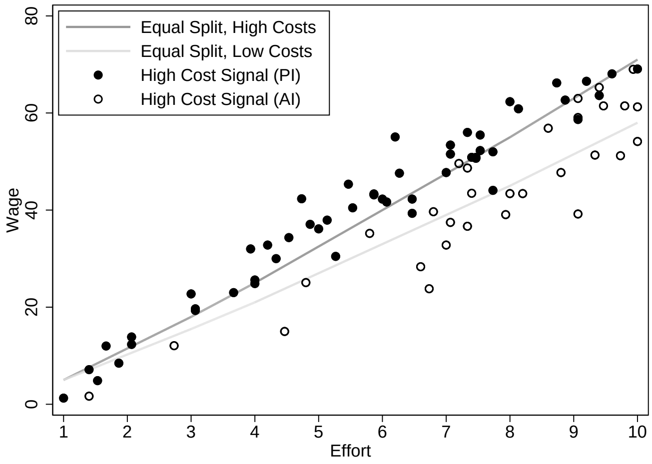


Figure A.3: Wage-Effort Relationship under High Cost Signals (PI and AI)



Note: One data point in the above plot represents the vector of average wage and average effort chosen in one relationship between a principal and an agent, where the averages are taken over all periods. “Equal Split, High Costs” (“Equal Split, Low Costs”) indicates the wage which splits the surplus equally between the parties for every given effort under the assumption of high (low) costs, where linear interpolation is applied.

Figure A.4: Effort in Case of Low Costs (AC and AI)

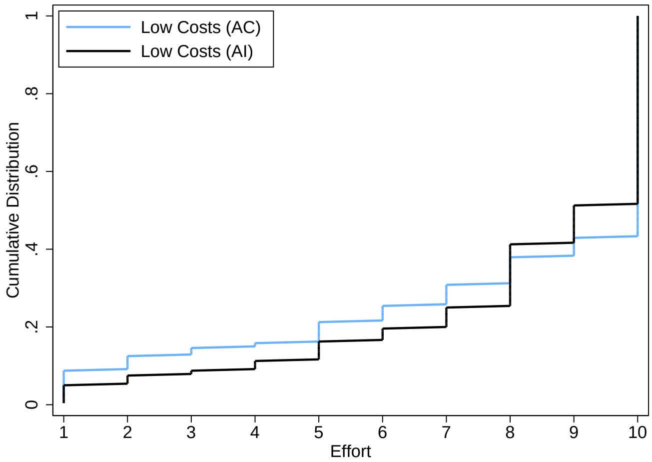


Figure A.5: Wage over Time (PC and PI)

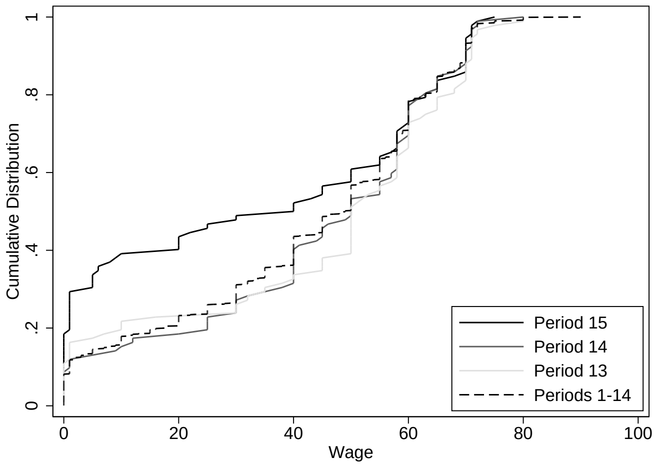
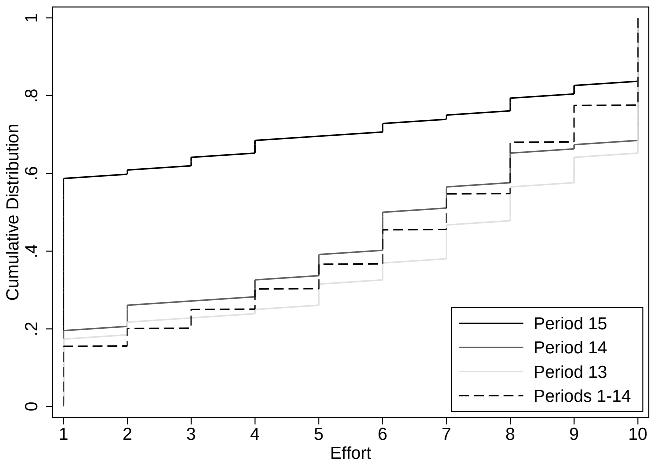


Figure A.6: Effort over Time (PC and PI)



A.2 Formal Definitions and Proofs

A.2.1 Complete Information

We first formulate the definitions of the strategy profiles in the complete information game more precisely, taking into account that the stage games are extensive-form games, so that principals and agents act after different histories.

Let $h_{-1}^P = \varnothing$ and $h_t^P = ((w_0, e_0), ..., (w_t, e_t))$ for $t \geq 0$. A strategy σ_P of the principal ($i = P$) maps every history h_{t-1}^P for $t \in \{0, 1, 2, ...\}$ into an action w_t. Next, let $h_0^A = w_0$ and $h_t^A = ((w_0, e_0), ..., (w_{t-1}, e_{t-1}), w_t)$ for $t \geq 1$. A strategy σ_θ of the agent of type θ maps every history h_t^A for $t \in \{0, 1, 2, ...\}$ into an action e_t.

Definition A.1. *(i) A trigger strategy profile (TP) of the RPA game with complete information and type $\theta \in \{L, H\}$ is given by a wage-effort vector (w^*, e^*) and pure strategies σ_i for $i \in \{P, \theta\}$ with*

$$\begin{aligned}
\sigma_P\left(h_{t-1}^P\right) &= \begin{cases} w^* & \textit{if} \quad h_{t-1}^P = \varnothing \textit{ or } h_{t-1}^P = ((w^*, e^*), ..., (w^*, e^*)), \\ 0 & \textit{otherwise;} \end{cases} \\
\sigma_\theta\left(h_t^A\right) &= \begin{cases} e^* & \textit{if} \quad h_t^A = w^* \textit{ or } h_t^A = ((w^*, e^*), ..., (w^*, e^*), w^*), \\ 1 & \textit{otherwise.} \end{cases}
\end{aligned}$$

(ii) A forgiving cut-off strategy profile (FCP) of the RPA game with complete information and type $\theta \in \{L, H\}$ is given by a wage-effort vector (w^*, e^*) and pure strategies σ_i for $i \in \{P, \theta\}$ with

$$\begin{aligned}
\sigma_P\left(h_{t-1}^P\right) &= \begin{cases} 0 & \textit{if} \quad \exists t' \leq t-1 \textit{ s.t. } w_{t'} \geq w^* \textit{ and } e_\tau < e^* \textit{ for all } \tau \in \{t', ..., t-1\}, \\ w^* & \textit{otherwise;} \end{cases} \\
\sigma_\theta\left(h_t^A\right) &= \begin{cases} 1 & \textit{if} \quad \begin{cases} w_{t'} < w^* \textit{ for all } t' \in \{0, \ldots, t\}, \\ \exists t' \textit{ where } 1 \leq t' \leq t \textit{ s.t., for all } \tau \in \{t', \ldots, t\}, w_\tau < w^* \textit{ and } e_{t'-1} \geq e^*, \end{cases} \\ e^* & \textit{otherwise.} \end{cases}
\end{aligned}$$

Proof of Proposition 2.1. We confine ourselves to the case of trigger-strategy profiles.[1] Fix a TP sustaining (w^*, e^*). We start by considering one-shot deviations of the principal. First, consider any history without deviation of either party. Then the resulting subgame involves only wages w^* and efforts e^*. The normalized discounted sum of equilibrium payoffs of the principal is then $S(e^*) + S(1) - w^*$. The most profitable downward deviation is to choose $w = 0$, yielding $S(1)$. The deviation is thus not profitable if and only if Condition (2.1) of Proposition 2.1 holds. Second, consider any history where a player has deviated. Then, in the corresponding subgame equilibrium outcome, the principal pays a zero wage forever and has an expected average payoff of $S(1)$. Any deviation to a positive wage is

[1] For the proof concerning the case of FCP, see Appendix A.3.1.

costly, without having a positive effect on the effort of the agent. Thus, deviation is not profitable.

Next, consider one-shot deviations of the agent. First, consider any history without deviation of either party. The normalized discounted sum of equilibrium payoffs of the agent is then $w^* - C(e^*, \theta)$. The most profitable downward deviation is to choose minimal efforts, yielding normalized discounted payoffs $(1-\delta) w^*$. The deviation is thus not profitable if and only if Condition (2.2) of Proposition 2.1 holds. Second, consider any history where some player has previously deviated. The agent's normalized discounted payoffs from following the equilibrium strategy is thus zero. Deviation to any effort above zero would only reduce the instantaneous payoffs without generating higher continuation payoffs. Thus, deviation is not profitable. □

A.2.2 Incomplete Information

Pooling Equilibria

The pooling strategy profiles described in Section 2.5.2 are straightforward adaptations of the corresponding strategy profiles under complete information. We confine ourselves to a detailed treatment of pooling trigger strategy equilibria; the analysis for the pooling forgiving cut-off equilibria is similar.[2] The adaptation of the forgiving cut-off strategy is analogous.

Definition A.2. *Denote the beliefs of the principal after a history h_{t-1}^P as $\mu_t\left(h_{t-1}^P\right)$. A pooling trigger strategy profile (PTP) of the RPA game with incomplete information is given by a wage-effort vector (w^*, e^*) and pure strategies σ_i for $i \in \{P, L, H\}$ with*

$$\begin{aligned}
\sigma_P\left(h_{t-1}^P, \mu_t\left(h_{t-1}^P\right)\right) &= \begin{cases} w^* & \text{if} \quad h_{t-1}^P = \varnothing \text{ or } h_{t-1}^P = ((w^*, e^*), ..., (w^*, e^*)) \\ 0 & \text{otherwise,} \end{cases} \\
\text{for } \theta \in \{L, H\}, \sigma_\theta\left(h_t^A\right) &= \begin{cases} e^* & \text{if} \quad h_t^A = w^* \text{ or } h_t^A = ((w^*, e^*), ..., (w^*, e^*), w^*) \\ 1 & \text{otherwise.} \end{cases}
\end{aligned}$$

Definition A.3. *A pooling trigger strategy equilibrium (PTE) is a PTP that forms a weak perfect Bayesian equilibrium of the RPA game with incomplete information, with beliefs $\mu_t\left(h_{t-1}^P\right) \equiv \mu$ for all $t \in \{0, 1, ...\}$ and all h_{t-1}^P.*

Proof of Proposition 2.2. We restrict attention to the PTE.[3] In a PTE, both types of players have the same strategy, corresponding to a TE sustaining (w^*, e^*). The non-deviation conditions of each type of agent are thus the same as for the corresponding agent under complete information (in the TE); thus they are given by (2.2). The more restrictive non-deviation condition is the one of the high type. Thus, the non-deviation condition of both types of agents in the PTE holds if and only if (2.4) does. The non-deviation conditions for the principal are the same as for the TE, as they also expect an

[2] We provide the details in the Appendix A.3.2.

[3] For the proof of the pooling forgiving cut-off equilibrium, see Appendix A.3.2.

effort of e^* if and only if they choose w^*. Finally, as the strategies are not informative about agents' types, it is consistent with Bayes' law to stick to prior beliefs. $\square$

Separating Equilibria

We first define our notion of separating equilibrium more precisely. Very complicated separating equilibria are conceivable in principle, but we confine ourselves to the simple separating trigger-strategy equilibrium mentioned in the text.

Definition A.4. *A separating trigger-strategy equilibrium (STE) in the RPA game with incomplete information is a weak perfect Bayesian equilibrium satisfying the following properties:*
(i) Initial wages are $w^* \geq 0$.
(ii) There exist $e^H < e^L$ and $w^H < w^L$ and strategies of the principal, σ_P, and agents, σ_θ, where $\theta \in \{L, H\}$, such that

$$(a)\, \sigma_\theta(h_t^A) = \begin{cases} e^\theta & \text{if } h_t^A = w^* \text{ or } h_t^A = ((w^*, e^\theta), (w^\theta, e^\theta), ..., (w^\theta, e^\theta), w^\theta), \\ e^H & \text{if } \theta = L \text{ and } h_t^A = ((w^*, e^H), (w^H, e^H), ..., (w^H, e^H), w^H), \\ 1 & \text{otherwise.} \end{cases}$$

$$(b)\, \sigma_P(h_{t-1}^P, \mu_t(h_{t-1}^P)) = \begin{cases} w^* & \text{if } h_{t-1}^P = \emptyset, \\ w^\theta & \text{if } h_{t-1}^P = (w^*, e^\theta) \text{ or } ((w^*, e^\theta), (w^\theta, e^\theta), ..., (w^\theta, e^\theta)), \\ 0 & \text{otherwise,} \end{cases}$$

(iii) Beliefs of the principal correspond to prior beliefs μ in period one. In periods two and following, the principal who has observed an agent playing e^θ in all previous periods believes he is of type θ with probability one. In all other cases, the principal's beliefs are arbitrary.

Proof of Proposition 2.3. For the proof, we have to show that the following holds. First, the principal is *sequentially rational*: If she faces an agent whose strategy is described by (ii)(a) in Definition A.4 the strategy described by (i) and (ii)(b) maximizes her expected payoff. Second, the agent is *sequentially rational*: If he faces a principal whose strategy is described by (i) and (ii)(b) in Definition A.4, the strategy described by (ii)(a) maximizes his expected payoff. Third, the principal's beliefs are consistent with Bayes' Law wherever applicable. We address the two sequential rationality requirements in two lemmas.

Lemma A.1. *Given the strategy of the principal, the agent's behavior is sequentially rational if the following conditions hold:*

$$c\left(e^L, H\right) - c\left(e^H, H\right) \geq \delta\left(w^L - w^H\right) \tag{ICH}$$

$$\delta\left(w^L - w^H\right) \geq c\left(e^L, L\right) - c\left(e^H, L\right) \tag{ICL}$$

$$\delta w^H \geq c\left(e^H, H\right) \tag{PCH}$$

Proof of Lemma A.1. (i) As long as neither player has previously deviated from the equilibrium, there are two possible types of deviations of the agent. First, there are deviations that trigger minimal wages forever, namely all deviations in periods 2 and following and all first-period deviations in period 1 to $e \notin \{e^L, e^H\}$. Second, the agent of type θ can deviate in period 1 by choosing $e^{\widetilde{\theta}}$ with $\widetilde{\theta} \neq \theta$.

To avoid the first type of deviation, we require participation constraints: Agents must be better off by playing the proposed equilibrium for their own type rather than triggering minimal wages forever and therefore minimal continuation profits. The optimal deviation is to choose minimal efforts. If such deviation takes place in periods 2,3,..., this yields a normalized discounted sum of payoffs of $(1-\delta)\, w^\theta$, whereas equilibrium play yields $w^\theta - C\left(e^\theta, \theta\right)$. Thus, sequential rationality of type H agents requires that PCH holds and, in addition, sequential rationality of type L agents requires $\delta w_L \geq c\left(e^L, L\right)$. Given (PCH), the latter inequality holds if (ICL) does. Similar arguments show that, in period 1, it is not profitable to deviate to any $e \notin \{e^L, e^H\}$, as this would also trigger minimal wages forever.

To avoid the second type of deviation, we require incentive constraints: Agents must be better off by choosing the proposed equilibrium action for their own type in period 1 rather than the action for the other type. For type L, the normalized discounted sum of payoffs after such a one-shot deviation is $(1-\delta)(w^* - c(e^H, L)) + \delta(w^H - c(e^H, L))$, whereas it is $(1-\delta)(w^* - c(e^L, L)) + \delta(w^L - c(e^L, L))$ in equilibrium. These deviations are not profitable if and only if (ICL) holds.

Next consider an agent of type H: If this player deviates to e^L in period 0, he will choose minimal effort ever after, triggering zero wages. Thus, such a deviation is not profitable if and only if

$$(1-\delta)\left(w^* - c\left(e^H, H\right)\right) + \delta\left(w^H - c\left(e^H, H\right)\right) \geq (1-\delta)\left(w^* - c\left(e^L, H\right)\right) + \delta\,(1-\delta)\, w^L$$

or equivalently

$$(1-\delta)\, c\left(e^L, H\right) - c\left(e^H, H\right) \geq \delta\,(1-\delta)\, w^L - \delta w^H.$$

This condition follows from (ICH) and (PCH) after simple manipulations.

(ii) Next, suppose the agent of type L has deviated to e^H in period 1 (and then followed the equilibrium prescription of continuing to play e^H), but no other deviation has occured. The normalized discounted sum of payoffs of the agent in such a subgame is $w^H - c(e^H, L)$. Any other choice of effort triggers zero wages and thus a normalized discounted sum of payoffs $(1-\delta)\, w^H$. Thus, the non-deviation condition for type L is $\delta w^H \geq c(e^H, L)$. But (PCH) implies $\delta w^H \geq c(e^H, H)$. Hence, such deviations are not profitable for the agent of type L.

(iii) Finally, consider any history in which some other deviation has taken place. Then, the continuation payoff of the agent is zero, independent of his effort choices. Thus, the equilibrium prescription of supplying zero effort is a (weakly) best response. □

Lemma A.2. *Given the strategy of the agent, the principal's proposed strategy is sequentially rational if and only if*

$$S(e^H) - w^H \geq 0$$
$$S(e^L) - w^L \geq 0$$
$$pS\left(e^L\right) + (1-p)\,S\left(e^H\right) - (1-\delta)\,w^* - \delta\left[pw^L + (1-p)\left(w^H\right)\right] \geq 0.$$

Proof of Lemma A.2. After any history in which at least one player has deviated, the principal obtains minimal profits no matter whether she pursues the equilibrium strategy or deviates. Thus, we only need to consider histories for which no player has deviated. The left-hand sides in the three inequalities in the lemma are the average expected equilibrium profits in any such history. Deviation leads to minimal average expected profits. Hence deviation is not profitable if the three inequalities in the lemma hold. □

The least costly way for the principal to guarantee that choosing e^θ is sequentially rational for the agent of type $\theta \in \{H, L\}$ is to set $w^* = 0$ and to choose w^L and w^H so that $\delta w^L = c\left(e^L, L\right) - c\left(e^H, L\right) + c\left(e^H, H\right)$ and $\delta w^H = c\left(e^H, H\right)$, so that (ICL) and (PCH) hold with equality. These equalities also imply (PCL). With these wage choices, the sequential rationality conditions from Lemma 2 become

$$\delta S(e^H) \geq c\left(e^H, H\right) \tag{A.1}$$
$$\delta S(e^L) \geq c\left(e^L, L\right) - c\left(e^H, L\right) + c\left(e^H, H\right) \tag{A.2}$$
$$pS\left(e^L\right) + (1-p)\,S\left(e^H\right) \geq p\left(c\left(e^L, L\right) - c\left(e^H, L\right) + c\left(e^H, H\right)\right) + (1-p)\,c\left(e^H, H\right) \tag{A.3}$$

(A.1) and (A.2) are identical with (2.5) and (2.6). Moreover, they imply (A.3).

This completes the proof of sequential rationality. Consistency with Bayes' rule is immediate: In period 0, the principal correctly applies her prior beliefs. In any history where the agent has always played e^θ, the posterior probability of type θ in the proposed equilibrium is 1. In any other history, Bayes' Law cannot be used to calculate posteriors, as such histories do not arise in equilibrium. □

A.3 Technical Appendix

In section A.3.1, we provide the proof of Proposition 2.1 for the case of FCE in the RPA game. In section A.3.2, we turn to the case of incomplete information and provide a proof of Proposition 2.2 for the case of PFCE. In Section, A.3.3, we address the infinitely *repeated agent-principal (RAP) game.* This game is identical to the RPA game, except that the agent moves before the principal in each stage game. The behavioural predictions obtained for this game therefore relate to behaviour observed in our treatments AC and AI.

A.3.1 Proof of Proposition 2.1 in Case of FCE

For convenience, we restate Proposition 2.1 for the case of FCE.

Proposition 2.1 (FCE). Consider the RPA game with complete information and cost type $\theta \in \{L, H\}$. Then, a wage-effort vector (w^*, e^*) is *sustainable as an FCE for* δ *and* θ if and only if

$$S\left(e^*\right) \geq w^*, \tag{A.4}$$

$$w^* \geq \frac{c\left(e^*, \theta\right)}{\delta}. \tag{A.5}$$

Proof of Proposition 2.1 (FCE). Fix an FCP sustaining (w^*, e^*). We start by considering one-shot deviations of the principal. First, consider any history where there has been no previous deviation of either party, or this deviation occured sufficiently long ago that the outcome path of the corresponding subgame equilibrium involves only wages w^* and efforts e^*.[4] The normalized discounted sum of payoffs for the principal is then $S\left(e^*\right) + S(1) - w^*$. The most profitable downward deviation is to choose $w = 0$, yielding a normalized discounted sum of payoffs of $(1-\delta)\, S(1) + \delta\left(S\left(e^*\right) + S(1) - w^*\right) = S(1) + \delta\left(S\left(e^*\right) - w^*\right)$ The deviation is thus not profitable if and only if $(1-\delta)\left(S\left(e^*\right) - w^*\right) \geq 0$, which is equivalent with Condition (A.4). Second, consider any history where the agent has triggered a deviation from which he has not returned. Then, in the corresponding subgame equilibrium outcome, the principal pays a zero wage at the corresponding node, whereas the agent chooses e^* thereafter. In later periods, the players choose e^* and w^*, respectively. The normalized discounted sum of subgame equilibrium payoffs for the principal is thus $(1-\delta)\left(S\left(e^*\right) + S(1)\right) + \delta\left(S\left(e^*\right) + S(1) - w^*\right) = S\left(e^*\right) + S(1) - \delta w^*$. Deviation to any wage above zero would only reduce the instantaneous payoffs without generating higher continuation payoffs; thus such a deviation is not profitable.

Next, consider one-shot deviations of the agent. First, consider any history where there has been no previous deviation of either party, or this deviation occured sufficiently long ago that the outcome path at the corresponding subgame involves only wages w^*

[4] In particular, this case comprises any history where the principal has triggered a deviation from which she has not returned: Then, the agent has already responded by choosing minimal effort before the principal moves again and the future equilibrium behavior of the principal (agent) is to choose w^* (e^*).

and efforts e^*.[5] The normalized discounted sum of subgame equilibrium payoffs of the agent is then $w^* - C(e^*,\theta)$. The most profitable downward deviation is to choose minimal efforts, so that the normalized discounted sum of subgame equilibrium payoffs is $(1-\delta)w^* + \delta((1-\delta)(-C(e^*,\theta)) + \delta(w^* - C(e^*,\theta))) = (1-\delta+\delta^2)w^* - \delta C(e^*,\theta)$. The deviation is thus not profitable if $\delta(1-\delta)w^* \geq (1-\delta)C(e^*,\theta)$, which is equivalent with Condition (A.5). Second, consider any history where the principal has previously triggered a deviation from which she has not returned. Then, in the corresponding subgame equilibrium outcome, the agent exerts minimal effort at the corresponding node. In later periods, the players choose e^* and w^*, respectively. The agent's normalized discounted sum of subgame equilibrium payoffs is thus $(1-\delta)w^* + \delta(w^* - C(e^*,\theta)) = w^* - \delta C(e^*,\theta)$. Deviation to any effort above zero would only reduce the instantaneous payoffs without generating higher continuation payoffs; thus the deviation is not profitable. □

A.3.2 Proof of Proposition 2.2 in Case of PFCE

For convenience, we restate Proposition 2.2 for the case of PFCE.

Proposition 2.2 (PFCE). Consider the RPA game with incomplete information. Then, the wage-effort vector (w^*, e^*) is sustainable as a PFCE given δ if and only if

$$S(e^*) \geq w^* \tag{A.6}$$

$$w^* \geq \frac{c(e^*, H)}{\delta}. \tag{A.7}$$

Proof of Proposition 2.2 (PFCE). In a PFCE, both types of players have the same strategy, corresponding to an FCE sustaining (w^*, e^*). The non-deviation conditions of each type of agent are thus the same as for the corresponding agent under complete information (in the FCE). The more restrictive non-deviation condition is the one of the high type. Thus, the non-deviation condition of both types of agents in the PFCE holds if and only if (A.7) does. The non-deviation condition for the principal is the same as for the FCE, as they also expect an effort of e^* if and only if they choose w^*. Thus, this condition is given by (A.6). □

A.3.3 The RAP Game

We now sketch the analysis for the RAP game. To define trigger strategy profiles and forgiving cut-off strategies in the case of complete information and pooling under incomplete information, only mild and obvious adjustments are necessary. Once these adjustments are in place, the following result is straightforward to prove.

Proposition A.1. *(i) Consider the RAP game with complete information with cost type $\theta \in \{L, H\}$. Then, a wage-effort vector (w^*, e^*) is sustainable as a TE given δ and θ if*

[5] In particular, this case comprises any history where the agent has triggered a deviation from which she has not returned (see the argument in the previous footnote).

and only if

$$\delta S(e^*) \geq w^* \tag{A.8}$$
$$w^* \geq c(e^*, \theta). \tag{A.9}$$

(ii) Consider the RAP game with incomplete information. Then, a wage-effort vector (w^*, e^*) *is sustainable as a PTE given* δ *if and only if*

$$\delta S(e^*) \geq w^* \tag{A.10}$$
$$w^* \geq c(e^*, H). \tag{A.11}$$

Proof of Proposition A.1. We prove only part (i), as the arguments for the pooling equilibrium mirror those for the RPA game. We start by considering one-shot deviations of the principal. First, consider any history without previous deviation of either party. Thus the outcome path in the corresponding subgame involves only wages w^* and efforts e^*. The normalized discounted equilibrium payoff sum of the principal is then $S(e^*) + S(1) - w^*$. The most profitable downward deviation is to choose $w = 0$, yielding a normalized discounted payoff sum of $S(1) + (1-\delta) S(e^*)$. The deviation is thus not profitable if and only if (A.8) holds. Second, consider any history with at least one deviation. Then, in the corresponding subgame equilibrium outcome, the principal pays a zero wage forever and obtains a normalized discounted payoff sum of $S(1)$. Deviation to any wage above zero would only reduce the instantaneous payoffs without generating higher continuation payoffs; thus such a deviation is not profitable.

Next, consider one-shot deviations of the agent. First, consider any history such that neither party has deviated so that the corresponding subgame involves only wages w^* and efforts e^*. The normalized discounted payoff sum of the agent is then $w^* - C(e^*,\theta)$. The most profitable downward deviation is to choose minimal efforts, yielding a normalized discounted payoff sum of 0. The deviation is thus not profitable if Condition (A.9) holds. Second, consider any history with at least one deviation. Then, in the corresponding subgame equilibrium outcome, the agent exerts minimal effort forever and obtains a normalized discounted payoff sum of zero. Deviation to any non-minimal effort would only reduce the instantaneous payoffs without generating higher continuation payoffs. Thus, there is no profitable deviation of the agent. □

Thus, in the case of complete information and incomplete information with pooling, the set of sustainable wage-effort vectors for the RAP game lies to the right of the corresponding set for the standard-order case in Figure 2.6. For any given wage, higher efforts are sustainable. The plot of the wage-effort relationship presented in Figure A.3 is consistent with this prediction. However, simple rearrangements of the equilibrium conditions in Proposition A.1 (i) show that the set of sustainable efforts remains unchanged by the order of the game.

It is equally straightforward to adapt the definition of separating trigger strategy profiles to the RAP game. The obvious difference is that agents move first, so that they can separate in the first period. Thus, a separating trigger strategy profile induces constant outcome paths, with efforts $e^H < e^L$ and wages $w^H < w^L$ such that (i) the type θ agents plays e^θ as long as no deviation from w^θ and e^θ has occured and (ii) the principal plays w^θ as long as no deviation from w^θ and e^θ has occured. With this definition, we obtain the following result.

Proposition A.2. *Consider the RAP game with incomplete information. Then, an STE sustaining e_H and e_L, exists for suitable wages w_H and w_L if and only if*

$$\delta S(e^H) \geq c\left(e^H, H\right), \tag{A.12}$$

$$\delta S(e^L) \geq c\left(e^L, L\right) - c\left(e^H, L\right) + c\left(e^H, H\right). \tag{A.13}$$

Proof of Proposition A.2. A separating trigger strategy profile in the RAP game induces constant outcome paths, with efforts $e^H < e^L$ and wages $w^H < w^L$ such that (i) the type θ agent chooses e^θ as long as no deviation from w^θ, e^θ has occured; type L chooses e^H after any history where he has always chosen e^H and the principal has chosen w^* (in period 1) and w^H (in periods 2, 3, ...). Both types of agents choose minimal effort otherwise; (ii) the principal plays w^θ as long as no deviation from w^θ, e^θ has occured and she choose a zero wage otherwise. We thus have to show that the following three conditions hold. First, we require sequential rationality of the principal. Second, we require sequential rationality of the agent. Third, the principal's belief are consistent with Bayes' rule wherever applicable. As the latter is obvious (see the corresponding proof for the RPA game), we only have to address the two sequential rationality requirements in the following lemmas.

Lemma A.3. *Consider the RAP game with incomplete information. Given the strategy of the principal in an STE, the proposed strategy of the agent is sequentially rational if*

$$c(e_L, H) - c(e_H, H) \geq w_L - w_H \qquad \textit{(ICAH)}$$

$$w_L - w_H \geq c(e_L, L) - c(e_H, L) \qquad \textit{(ICAL)}$$

$$w_H \geq c(e_H, H) \qquad \textit{(PCAH)}$$

Proof of Lemma A.3. (i) As long as neither player has previously deviated from the equilibrium, there are two possible types of deviations of the agent. First, there are deviations that trigger minimal wages forever, namely all deviations in periods 2 and following and all first-period deviations in period 1 to $e \notin \left\{e^L, e^H\right\}$. Second, the agent of type θ can deviate in period 1 by choosing $e^{\widetilde{\theta}}$ with $\widetilde{\theta} \neq \theta$.

To avoid the first type of deviation, we require participation constraints: Agents must be better off by playing the proposed equilibrium for their own type rather than by triggering minimal wages forever and therefore minimal continuation profits. The optimal deviation is to choose minimal efforts, which yields a normalized discounted sum of payoffs of 0, whereas equilibrium play yields $w^\theta - c\left(e^\theta, \theta\right)$. Thus, sequential rationality of type-H agents requires that PCAH holds. Similar arguments show that, in period 1, it is not profitable for type H to deviate to any $e \notin \{e^L, e^H\}$, as this would also trigger minimal wages forever. For type-L agents, sequential rationality requires that $w_L \geq c(e_L, L)$, which is implied by (ICAL) and (PCH).

To avoid the second type of deviation, we require incentive constraints: Agents must be better off by choosing the proposed equilibrium action for their own type in period 1 rather than the action for the other type. For type L, the normalized discounted sum of payoffs after such a one-shot deviation is $(1-\delta)(w^* - c(e^H, L)) + \delta(w^H - c(e^H, L))$, whereas it is $(1-\delta)\left(w^* - c\left(e^L, L\right)\right) + \delta\left(w^L - c\left(e^L, L\right)\right)$ in equilibrium. These deviations are not profitable if (ICAL) holds.

Next consider an agent of type H: if this player deviates to e^L in period 0, she will choose minimal effort ever after, triggering zero wages. Thus, such a deviation is not profitable if and only if

$$(1-\delta)\left(w^H - c\left(e^H, H\right)\right) + \delta\left(w^H - c\left(e^H, H\right)\right) \geq (1-\delta)\left(w^L - c\left(e^L, H\right)\right).$$

This condition follows from multiplying (ICAH) with $(1-\delta)$ and (PCAH) with δ and then adding up.

(ii) Next, suppose the agent of type L has deviated to e^H in period 1 (and then followed the equilibrium prescription of continuing to play e^H), but no other deviation has occured. The normalized discounted sum of payoffs of the agent in such a subgame is $w^H - c(e^H, L)$. Any other choice of effort triggers zero wages and thus a normalized discounted sum of payoffs 0. By (PCAH) and the definition of the cost function, $w^H \geq c(e^H, H) \geq c(e^H, L)$. Thus, such deviations are not profitable for type H.

(iii) Finally, consider any history in which some other deviation has taken place. Then, the continuation payoff of the agent is zero, independent of his effort choices. Thus, the equilibrium prescription of supplying zero effort is a (weakly) best response. □

Lemma A.4. *Consider the RAP game with incomplete information. Given the strategy of the agent in an STE, the proposed strategy of the principal is sequentially rational if and only if*

$$\begin{array}{rcll} \delta S(e_H) & \geq & w_H & \quad \textit{(PCPH)} \\ \delta S(e_L) & \geq & w_L & \quad \textit{(PCPL)} \end{array}$$

Proof of Lemma A.4. In any history such that one player has already deviated from the equilibrium, the strategy of the agent implies that the principal obtains a payoff of zero. Thus, suppose that no player has previously deviated from the equilibrium play, so that the principal observed e_θ in all previous periods. Then, the principal believes that the agent is of type θ with probability one. Thus, if the agent follows the equilibrium strategy in all subsequent periods, the normalized expected discounted sum of equilibrium payoffs is $S(e_\theta) - w_\theta$ for the principal. If she chooses the optimal deviation and sets a zero wage, the normalized discounted sum of payoffs is $(1-\delta)\,S(e_\theta)$. Thus, the principal is sequentially rational. $\square$

We now use the Lemmas to complete the proof of Proposition A.2. Let $w_H = c\,(e_H, H)$. Then (PCAH) holds. With $w_L = c\,(e_L, L) - c\,(e_H, L) + c\,(e_H, H)$, (ICAL) holds as well. As $w_L - w_H = c\,(e_L, L) - c\,(e_H, L)$, the assumption that $c\,(e_L, H) - c\,(e_H, H) > c\,(e_L, L) - c\,(e_H, L)$ implies (ICAH). Thus, by Lemma A.3 the agent is sequentially rational. Moreover, clearly for the proposed wages, conditions (PCPH) and (PCPL) hold, if (A.12) and (A.13) do. Thus, conditions (A.12) and (A.13) are also necessary for an STE: (PCAH) and (PCPH) imply (A.12); (PCAH), (ICAL) and (PCPL) imply (A.13). $\square$

Thus, the conditions for the existence are exactly the same as in the RPA game.

A.4 Instructions

A.4.1 Instructions for Buyers in Treatment PC

Allgemeine Erklärungen für Käufer

Sie nehmen nun an einer wirtschaftswissenschaftlichen Studie teil, die von diversen Forschungsförderungsstellen finanziert wird. Sie können dabei - abhängig von Ihren Entscheidungen – Geld verdienen. Es ist daher sehr wichtig, dass Sie diese Erklärungen genau durchlesen.

Diese Instruktionen dienen ausschliesslich Ihrer privaten Information. **Während der Studie herrscht ein absolutes Kommunikationsverbot.** Wenn Sie Fragen haben, dann richten Sie diese bitte an uns. Die Nichtbeachtung dieser Regel führt zum Ausschluss aus der Studie und von allen Zahlungen.

Zu Beginn der Studie erhalten Sie ein Startgeld von 10 Franken. Während der Studie sprechen wir nicht von Franken, sondern von Punkten. Im Verlauf der Studie können Sie einen weiteren Geldbetrag verdienen, indem Sie Punkte erzielen. Ihr gesamtes Einkommen wird also zunächst in Punkten berechnet. Die von Ihnen während der Studie erzielte Gesamtpunktezahl wird dann am Ende in Franken umgerechnet, dabei gilt

10 Punkte = 1 Franken.

Sollten Sie im Laufe der Studie Verluste machen, werden allfällige Verluste mit Ihrem Startgeld verrechnet. **Sie können jedoch Verluste immer durch eigene Entscheidungen <u>mit Sicherheit</u> ausschliessen!** Am Ende bekommen Sie von uns die während der Studie verdiente Punktezahl plus die 10 Franken Startgeld in **bar** ausbezahlt.

Zu Beginn der Studie wurden die TeilnehmerInnen zufällig in zwei Gruppen aufgeteilt: Käufer und Verkäufer. **Sie sind während der gesamten Studie ein Käufer.** Jeder Käufer wurde zufällig einem Verkäufer zugeordnet. Die Studie besteht aus 15 Runden. Sie sind in jeder der 15 Runden dem gleichen Verkäufer zugeordnet. Keiner der anderen Studienteilnehmer wird Ihre genaue Zuordnung oder Ihre Entscheidungen erfahren. Die Anonymität bleibt also gewahrt.

1

Kurzübersicht über die Studie

Zu Beginn der Studie wurde jedem Käufer zufällig ein Verkäufer zugeordnet. Die Studie umfasst 15 Runden, in denen die Zuordnung von Käufer und Verkäufer fixiert bleibt. In jeder Runde führt jeder Käufer mit seinem zugeordneten Verkäufer eine Transaktion durch: Der Käufer zahlt dem Verkäufer einen Preis für ein Produkt, dessen Qualität durch den Verkäufer bestimmt wird. Insbesondere erzielt der Käufer durch die Transaktion einen Gewinn, wenn er für das Produkt weniger bezahlt, als das Produkt ihm wert ist. Wie hoch der Wert des Produktes für den Käufer ist, hängt von der Qualität des Produktes ab. Der Verkäufer erzielt durch die Transaktion einen Gewinn, wenn er einen Preis erhält, der seine Produktionskosten übersteigt. Produktionskosten entstehen durch die Bereitstellung von Produktqualität. Eine höhere Qualität ist immer mit höheren Produktionskosten verbunden.

Bestimmung der Produktionskosten

Bevor Käufer und Verkäufer Transaktionen durchführen, werden die genauen Produktionskosten bestimmt. Ein Verkäufer kann entweder **hohe oder niedrige Produktionskosten** haben. Unter den hohen Produktionskosten kostet jede Qualitätsstufe mehr als unter den niedrigen Produktionskosten. Die eine Hälfte aller Verkäufer wird zufällig den hohen Produktionskosten zugewiesen, die andere Hälfte aller Verkäufer den niedrigen Produktionskosten zugeordnet. Diese Zuweisung bleibt über die gesamte Studie hinweg bestehen.

Ihr Verkäufer wird zu Beginn der Studie darüber informiert, ob er den hohen oder niedrigen Produktionskosten zugeordnet ist. Sie werden ebenfalls über die genauen Produktionskosten Ihres Verkäufers informiert.

Ablauf einer Transaktion

Nach Bestimmung der Produktionskosten führen Sie und Ihr zugeordneter Verkäufer in jeder der 15 Runden eine Transaktion durch. Der Ablauf ist folgendermassen organisiert:

1.Stufe: Sie bezahlen einen Preis und geben eine gewünschte Produktqualität an.

2.Stufe: Ihr Verkäufer erhält daraufhin den von Ihnen bezahlten Preis, wird über Ihre gewünschte Qualität informiert, und kann die Produktqualität wählen, die er an Sie liefern will. **Die von Ihnen gewünschte Qualität ist für die Wahl Ihres Verkäufers nicht bindend.**

3.Stufe: Sie erhalten die von Ihrem Verkäufer gewählte Produktqualität. Sie werden über Ihr Einkommen in der aktuellen Runde informiert. Danach beginnt die nächste Runde.

Die Einkommen aus allen 15 Runden werden am Ende des Experiments zusammengezählt, in Franken umgerechnet und zusammen mit dem Startgeld bar ausbezahlt.

Auf den nächsten Seiten beschreiben wir den Ablauf der Bestimmung der Produktionskosten und die Transaktion weiter im Detail.

2

Information über den genauen Ablauf der Studie

Bestimmung der Produktionskosten

Die Produktionskosten des Verkäufers hängen von der Produktqualität ab. Die Produktqualität kann nicht kleiner als 1 und nicht höher als 10 sein:

1 ≤ Produktqualität ≤ 10.

Die Produktionskosten sind umso höher, je höher die gewählte Qualität. **Die Produktionskosten hängen zudem davon ab, ob der Verkäufer den hohen oder niedrigen Produktionskosten zugewiesen ist.** Wenn Ihr Verkäufer den hohen Produktionskosten zugewiesen ist, dann sind seine mit jeder Qualitätsstufe verbundenen Produktionskosten höher als wenn er den niedrigen Produktionskosten zugeordnet ist. Die niedrigen und hohen Produktionskosten sind für jede Produktqualität in der nachstehenden Tabelle beschrieben. Die Tabelle sowie eine grafische Beschreibung finden Sie auch auf dem Zusatzblatt.

Produktqualität	1	2	3	4	5	6	7	8	9	10
Niedrige Produktionskosten	0	0.5	1	2	4	6	8	10	13	16
Hohe Produktionskosten	0	3	6	10	15	20	25	30	36	42

Der Computer weist zu Beginn der Studie per Zufall die eine Hälfte aller Verkäufer den hohen Produktionskosten und die andere Hälfte den niedrigen Produktionskosten zu.
Ihr Verkäufer wird auf dem Bildschirm über seine tatsächlichen Produktionskosten informiert. Die tatsächlichen Produktionskosten bleiben nach der einmaligen Zuordnung zu Beginn der ersten Periode in jeder der 15 Runden gleich. Sie werden ebenfalls über die tatsächlichen Produktionskosten Ihres Verkäufers informiert.
Im Fall der Zuordnung zu den niedrigen Kosten bekommen Sie folgende Information:

Der Computer hat per Zufall die eine Hälfte aller Verkäufer den hohen Produktionskosten und die andere Hälfte der Verkäufer den niedrigen Produktionskosten zugewiesen.

Ihr Verkäufer hat **niedrige** Produktionskosten.

3

Ihr Verkäufer wird folgendermassen über Ihre Produktionskosten informiert (im Beispiel sind die Produktionskosten wiederum niedrig):

> Der Computer hat per Zufall die eine Hälfte aller Verkäufer den hohen Produktionskosten und die andere Hälfte der Verkäufer den niedrigen Produktionskosten zugewiesen.
>
> Sie sind in dieser Studie den **niedrigen** Produktionskosten zugeordnet.

Wenn Ihr Verkäufer den hohen Kosten zugeordnet ist, sehen Sie und Ihr Verkäufer analoge Information auf Ihren Bildschirmen.

Sie und Ihr Verkäufer werden in jeder Runde, während sie Ihre Entscheidungen treffen, an die Ihrem Verkäufer zugewiesenen Produktionskosten erinnert.

Ablauf einer Transaktion

Nach Bestimmung der Produktionskosten führen Sie und Ihr zugeordneter Verkäufer in jeder der 15 Runden eine Transaktion durch. Der Ablauf jeder Runde ist wie folgt organisiert:

1.Stufe: Preisangabe und Qualitätswunsch.
2.Stufe: Festlegung der tatsächlichen Produktqualität.
3.Stufe: Bestimmung der Einkommen.

1.Stufe: Preisangabe und Qualitätswunsch

Zu Beginn jeder Runde bezahlt der Käufer einen Preis und gibt einen Qualitätswunsch an. **Die Käufer müssen folgende Regeln einhalten:**

- Der Preis darf nicht kleiner als 0 und nicht höher als 100 sein:

0 ≤ Preis≤ 100.

- Die gewünschte Produktqualität darf nicht kleiner als 1 und nicht höher als 10 sein:

1 ≤ gewünschte Produktqualität ≤ 10.

Die Wahl des Preises und die Angabe der gewünschten Qualität tätigen Sie über folgende Bildschirmanzeige:

Im obigen Beispiel der Bildschirmanzeige ist der Verkäufer den niedrigen Kosten zugewiesen. Sie sehen eine analoge Bildschirmanzeige, wenn Ihr Verkäufer hohe Kosten hat.

Klicken Sie auf den „OK“-Knopf um den im Feld eingegebenen Preis und Ihre eingegebene gewünschte Qualität definitiv zu machen. **Bitte beachten Sie hierbei stets die mit der gewünschten Qualität verbundenen Produktionskosten, die von der Zuteilung Ihres Verkäufers zu den hohen oder niedrigen Produktionskosten abhängen.** Solange Sie nicht „OK“ gedrückt haben, können Sie Ihre Wahl revidieren.

2.Stufe: Festlegung der tatsächlichen Produktqualität

Daraufhin erhält Ihr Verkäufer den bezahlten Preis und wird über Ihren Qualitätswunsch informiert. Dann kann Ihr Verkäufer bestimmen, welche Produktqualität er an Sie liefern will. **Die von Ihnen gewünschte Produktqualität ist für Ihren Verkäufer nicht bindend. Ihr Verkäufer kann exakt die von Ihnen gewünschte Qualität wählen, aber auch eine höhere oder tiefere Qualität.**
Die Qualität, die Ihr Verkäufer wählt, muss ein ganzzahliger Wert zwischen 1 und 10 sein:

1 ≤ tatsächliche Produktqualität ≤ 10.

5

Die Wahl der tatsächlichen Produktqualität tätigt Ihr Verkäufer über die folgende Bildschirmanzeige:

In obigem Beispiel der Bildschirmanzeige ist Ihr Verkäufer den niedrigen Kosten zugeordnet. Ihr Verkäufer erhält eine analoge Bildschirmanzeige, wenn er den hohen Kosten zugeordnet ist.

3.Stufe: Bestimmung der Einkommen

Das Einkommen Ihres Verkäufers hängt vom bezahlten Preis, der von Ihm gewählten Produktqualität und den Ihm zugewiesenen Produktionskosten ab.

Wenn Ihr Verkäufer den niedrigen Produktionskosten zugeordnet ist, dann berechnet sich sein Einkommen wie folgt:

Einkommen Ihres Verkäufers = Preis – niedrige Produktionskosten der gewählten Produktqualität

6

Wenn Ihr Verkäufer den hohen Produktionskosten zugeordnet ist, dann berechnet sich sein Einkommen wie folgt:

Einkommen Ihres Verkäufers = Preis – hohe Produktionskosten der gewählten Produktqualität

Das Einkommen Ihres Verkäufers ist demzufolge umso höher, je höher der von Ihnen angegebene Preis. Ausserdem gilt, dass das Einkommen Ihres Verkäufers umso höher ist, je tiefer seine Produktionskosten.
Wie oben beschrieben gilt, dass mehr Qualität immer mehr kostet. Darüber hinaus sind die Produktionskosten für jede gewählte Qualitätsstufe höher, wenn Ihr Verkäufer den hohen Produktionskosten zugewiesen ist, als wenn Ihr Verkäufer den niedrigen Produktionskosten zugeordnet ist. Die genauen Werte entnehmen Sie oben stehender Tabelle oder der Tabelle und Grafik auf dem Zusatzblatt.

Ihr Einkommen hängt davon ab, welchen Preis Sie bezahlt haben, und welchen Wert das Produkt für Sie hat. Der Wert des Produktes für Sie wird durch die von Ihrem Verkäufer gewählte tatsächliche Produktqualität wie folgt bestimmt.

Wert für Sie = 10* tatsächliche Produktqualität.

Eine tabellarische und grafische Beschreibung des Werts für den Käufer für jede tatsächliche Produktqualität entnehmen Sie dem Zusatzblatt.

Insgesamt ergibt sich Ihr Einkommen wie folgt:

Ihr Einkommen = Wert für Sie – Preis
= 10*tatsächliche Produktqualität - Preis

Ihr Einkommen ist folglich umso höher, je höher die von Ihrem Verkäufer gelieferte Produktqualität, da eine höhere Qualität einen höheren Wert für Sie ergibt. Gleichzeitig ist Ihr Einkommen umso höher, je tiefer der Preis, den Sie für das Produkt bezahlt haben.

Die Einkommen aller Verkäufer und Käufer werden in der oben beschriebenen Weise berechnet. Jeder Verkäufer kann also das Einkommen seines Käufers berechnen. Zudem kann jeder Käufer das Einkommen seines Verkäufers bestimmen, unter Berücksichtigung der Zuordnung seines Verkäufers zu den hohen oder niedrigen Produktionskosten.

Beachten Sie, dass Käufer und Verkäufer in jeder Periode auch Verluste erzielen können. Diese müssen aus dem jeweiligen Startgeld bzw. aus in anderen Perioden erzielten Einkommen bezahlt werden. Beachten Sie, dass eigene Verluste sowie Verluste für den anderen immer durch eigene Entscheidungen mit Sicherheit ausgeschlossen werden können.

7

Nachstehende Beispiele sollen die Bestimmung der Einkommen illustrieren.

Beispiel 1:
Ein Käufer bezahlt einen Preis von 35 und wünscht eine Qualität von 8. Sein Verkäufer wählt eine tatsächliche Qualität von 4.
Das Einkommen des Käufers ist: 4*10-35=5.
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen: 35-2=33.
Wenn der Verkäufer hohe Kosten hat, dann ist sein Einkommen: 35-10=25.

Beispiel 2:
Ein Käufer bezahlt einen Preis von 50 und wünscht eine Qualität von 10. Sein Verkäufer wählt eine tatsächliche Qualität von 8.
Das Einkommen des Käufers ist: 8*10-50=30.
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen: 50-10=40.
Wenn der Verkäufer hohe Kosten hat, dann ist sein Einkommen: 50-30=20.

Beispiel 3:
Ein Käufer bezahlt einen Preis von 25 und wünscht eine Qualität von 4. Sein Verkäufer wählt eine tatsächliche Qualität von 7.
Das Einkommen des Käufers ist: 7*10-25=45.
Wenn der Verkäufer hohe Kosten hat, dann ist sein Einkommen: 25-8=17.
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen: 25-25=0.

Am Ende jeder Runde wird Ihrem Verkäufer eine Zusammenfassung der aktuellen Runde auf dem Bildschirm präsentiert. Diese beinhaltet:

- Welchen Preis Sie angegeben haben.
- Ihre gewünschte Qualität.
- Die tatsächliche Qualität, die Ihr Verkäufer gewählt hat.
- Das Einkommen Ihres Verkäufers in dieser Runde.

Sie erhalten ebenfalls eine Zusammenfassung über:

- Ihren angegebenen Preis.
- Ihre gewünschte Qualität.
- Die tatsächliche Qualität, die Ihr Verkäufer gewählt hat.
- Ihr erzieltes Einkommen in dieser Runde.

Nachdem die Zusammenfassung zu sehen ist, ist die aktuelle Runde abgeschlossen. Danach beginnt die nächste Runde. Insgesamt besteht die Studie aus 15 Runden, und Sie sind in jeder der 15 Runden demselben Verkäufer zugeordnet.

Die Studie beginnt erst dann, wenn alle TeilnehmerInnen mit dem Ablauf der Studie vollständig vertraut sind. Um dies sicher zu stellen, bitten wir Sie unten stehende Übungsaufgaben zu lösen.

8

Übungsaufgaben

Bitte lösen Sie diese **Aufgaben vollständig und unter Angabe des Lösungswegs**. Wenn Sie Fragen haben, wenden Sie sich bitte an die Leiter des Experiments.

Aufgabe 1:
Der bezahlte Preis des Käufers ist 55 und seine gewünschte Qualität ist 9. Der zugewiesene Verkäufer wählt eine tatsächliche Qualität von 8.

Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Aufgabe 2:
Der bezahlte Preis des Käufers ist 60 und seine gewünschte Qualität ist 9. Der zugewiesene Verkäufer wählt eine tatsächliche Qualität von 5.

Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Aufgabe 3:
Der bezahlte Preis des Käufers ist 9 und seine gewünschte Qualität ist 4. Der zugewiesene Verkäufer wählt eine tatsächliche Qualität von 1.

Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Aufgabe 4:
Der bezahlte Preis des Käufers ist 40 und seine gewünschte Qualität ist 10. Der zugewiesene Verkäufer wählt eine tatsächliche Qualität von 8.

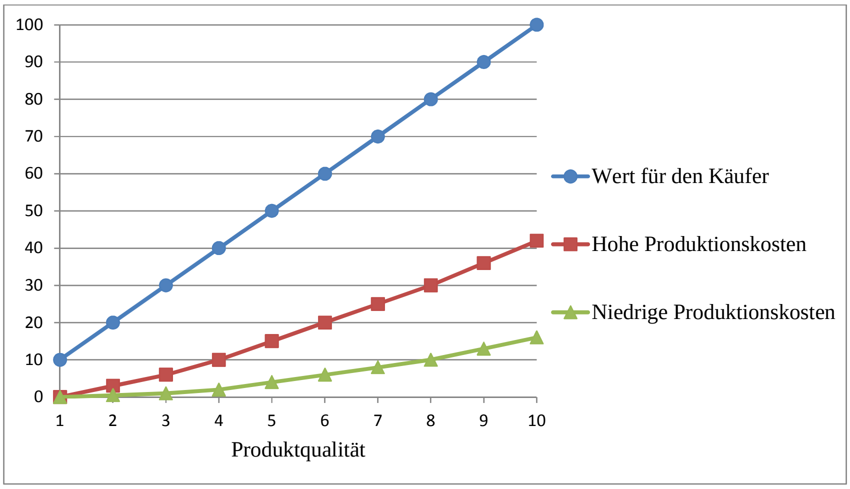
Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Wenn Sie mit dem Lösen der Übungsaufgaben fertig sind, empfehlen wir Ihnen, sich die Aufgaben und deren Lösungen noch einmal anzusehen. Anschliessend überlegen Sie sich bitte, welche Entscheidungen Sie im Experiment treffen wollen.

9

Zusatzblatt

Produktqualität	1	2	3	4	5	6	7	8	9	10
Niedrige Produktionskosten	0	0.5	1	2	4	6	8	10	13	16
Hohe Produktionskosten	0	3	6	10	15	20	25	30	36	42
Wert für den Käufer	10	20	30	40	50	60	70	80	90	100



10

A.4.2 Instructions for Sellers in Treatment PI

Allgemeine Erklärungen für Verkäufer

Sie nehmen nun an einer wirtschaftswissenschaftlichen Studie teil, die von diversen Forschungsförderungsstellen finanziert wird. Sie können dabei - abhängig von Ihren Entscheidungen – Geld verdienen. Es ist daher sehr wichtig, dass Sie diese Erklärungen genau durchlesen.

Diese Instruktionen dienen ausschliesslich Ihrer privaten Information. **Während der Studie herrscht ein absolutes Kommunikationsverbot.** Wenn Sie Fragen haben, dann richten Sie diese bitte an uns. Die Nichtbeachtung dieser Regel führt zum Ausschluss aus der Studie und von allen Zahlungen.

Zu Beginn der Studie erhalten Sie ein Startgeld von 10 Franken. Während der Studie sprechen wir nicht von Franken, sondern von Punkten. Im Verlauf der Studie können Sie einen weiteren Geldbetrag verdienen, indem Sie Punkte erzielen. Ihr gesamtes Einkommen wird also zunächst in Punkten berechnet. Die von Ihnen während der Studie erzielte Gesamtpunktezahl wird dann am Ende in Franken umgerechnet, dabei gilt

10 Punkte = 1 Franken.

Sollten Sie im Laufe der Studie Verluste machen, werden allfällige Verluste mit Ihrem Startgeld verrechnet. **Sie können jedoch Verluste immer durch eigene Entscheidungen mit Sicherheit ausschliessen!** Am Ende bekommen Sie von uns die während der Studie verdiente Punktezahl plus die 10 Franken Startgeld in **bar** ausbezahlt.

Zu Beginn der Studie wurden die TeilnehmerInnen zufällig in zwei Gruppen aufgeteilt: Käufer und Verkäufer. **Sie sind während der gesamten Studie ein Verkäufer.** Jeder Verkäufer wurde zufällig einem Käufer zugeordnet. Die Studie besteht aus 15 Runden. Sie sind in jeder der 15 Runden dem gleichen Käufer zugeordnet. Keiner der anderen Studienteilnehmer wird Ihre genaue Zuordnung oder Ihre Entscheidungen erfahren. Die Anonymität bleibt also gewahrt.

1

Kurzübersicht über die Studie

Zu Beginn der Studie wurde jedem Verkäufer zufällig ein Käufer zugeordnet. Die Studie umfasst 15 Runden, in denen die Zuordnung von Verkäufer und Käufer fixiert bleibt. In jeder Runde führt jeder Käufer mit seinem zugeordneten Verkäufer eine Transaktion durch: Der Käufer zahlt dem Verkäufer einen Preis für ein Produkt, dessen Qualität durch den Verkäufer bestimmt wird. Insbesondere erzielt der Käufer durch die Transaktion einen Gewinn, wenn er für das Produkt weniger bezahlt, als das Produkt ihm wert ist. Wie hoch der Wert des Produktes für den Käufer ist, hängt von der Qualität des Produktes ab. Der Verkäufer erzielt durch die Transaktion einen Gewinn, wenn er einen Preis erhält, der seine Produktionskosten übersteigt. Produktionskosten entstehen durch die Bereitstellung von Produktqualität. Eine höhere Qualität ist immer mit höheren Produktionskosten verbunden.

Bestimmung der Produktionskosten

Bevor Käufer und Verkäufer Transaktionen durchführen, werden die genauen Produktionskosten bestimmt. Ein Verkäufer kann entweder **hohe oder niedrige Produktionskosten** haben. Unter den hohen Produktionskosten kostet jede Qualitätsstufe mehr als unter den niedrigen Produktionskosten. Die eine Hälfte aller Verkäufer wird zufällig den hohen Produktionskosten zugewiesen, die andere Hälfte aller Verkäufer den niedrigen Produktionskosten zugeordnet. Diese Zuweisung bleibt über die gesamte Studie hinweg bestehen.

Sie werden zu Beginn der Studie darüber informiert, ob Sie den hohen oder niedrigen Produktionskosten zugeordnet sind. Der Ihnen zugeordnete Käufer ist nicht darüber informiert, ob Sie den hohen oder niedrigen Produktionskosten zugeordnet sind. Sie haben jedoch die Möglichkeit Ihrem Käufer zu Beginn der Studie einmalig eine Nachricht über Ihre Produktionskosten zu übermitteln.

Ablauf einer Transaktion

Nach Bestimmung der Produktionskosten führen Sie und Ihr zugeordneter Käufer in jeder der 15 Runden eine Transaktion durch. Der Ablauf ist folgendermassen organisiert:

1.Stufe: Ihr Käufer bezahlt einen Preis und gibt eine gewünschte Produktqualität an.

2.Stufe: Sie erhalten daraufhin den bezahlten Preis, werden über die gewünschte Qualität informiert, und können die Produktqualität wählen, die Sie an Ihren Käufer liefern wollen. **Die gewünschte Qualität Ihres Käufers ist für Ihre Wahl der tatsächlichen Qualität nicht bindend.**

3.Stufe: Ihr Käufer erhält die von Ihnen gewählte Produktqualität. Sie werden über Ihr Einkommen in der aktuellen Runde informiert. Danach beginnt die nächste Runde.

Die Einkommen aus allen 15 Runden werden am Ende des Experiments zusammengezählt, in Franken umgerechnet und zusammen mit dem Startgeld bar ausbezahlt.

Auf den nächsten Seiten beschreiben wir den Ablauf der Bestimmung der Produktionskosten und die Transaktion weiter im Detail.

2

Information über den genauen Ablauf der Studie

Bestimmung der Produktionskosten

Die Produktionskosten des Verkäufers hängen von der Produktqualität ab. Die Produktqualität kann nicht kleiner als 1 und nicht höher als 10 sein:

1 ≤ Produktqualität ≤ 10.

Die Produktionskosten sind umso höher, je höher die gewählte Qualität. **Die Produktionskosten hängen zudem davon ab, ob der Verkäufer den hohen oder niedrigen Produktionskosten zugewiesen ist.** Wenn Sie den hohen Produktionskosten zugewiesen sind, dann sind Ihre mit jeder Qualitätsstufe verbundenen Produktionskosten höher als wenn Sie den niedrigen Produktionskosten zugeordnet sind. Die niedrigen und hohen Produktionskosten sind für jede Produktqualität in der nachstehenden Tabelle beschrieben. Die Tabelle sowie eine grafische Beschreibung finden Sie auch auf dem Zusatzblatt.

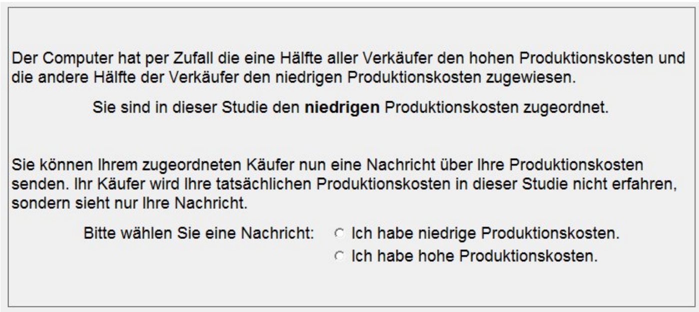
Produktqualität	1	2	3	4	5	6	7	8	9	10
Niedrige Produktionskosten	0	0.5	1	2	4	6	8	10	13	16
Hohe Produktionskosten	0	3	6	10	15	20	25	30	36	42

Der Computer weist zu Beginn der Studie per Zufall die eine Hälfte aller Verkäufer den hohen Produktionskosten und die andere Hälfte den niedrigen Produktionskosten zu.
Sie werden auf dem Bildschirm über Ihre tatsächlichen Produktionskosten informiert. Die tatsächlichen Produktionskosten bleiben nach der einmaligen Zuordnung zu Beginn der ersten Periode in jeder der 15 Runden gleich.
Ihrem Käufer ist nicht bekannt, welche Produktionskosten Ihnen tatsächlich zugeordnet sind, d.h. er weiss nicht, ob Sie hohe oder niedrige Produktionskosten haben.

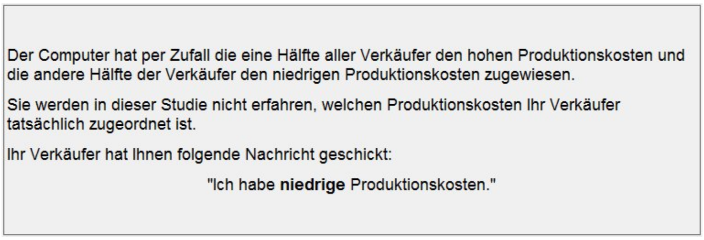
Sie können Ihrem Käufer jedoch zu Beginn der Studie einmalig eine Nachricht senden. Sie haben die Wahl zwischen der Nachricht „Ich habe niedrige Produktionskosten.“ und der Nachricht „Ich habe hohe Produktionskosten.“ **Unabhängig von Ihren tatsächlichen Kosten können Sie eine der beiden Nachrichten frei auswählen.**

Zum Beispiel, wenn Sie den niedrigen Kosten zugeordnet sind, dann sehen sie folgende Information auf Ihrem Bildschirm:

3



Wenn Sie den hohen Kosten zugeordnet sind, sehen Sie eine analoge Bildschirmanzeige.
Wählen Sie die von Ihnen gewünschte Nachricht aus. Solange Sie den „OK“-Knopf nicht angeklickt haben, können Sie Ihre Wahl noch verändern. Die von Ihnen gewählte Nachricht wird dann an Ihrem Käufer übermittelt und kann nicht mehr verändert werden. Zum Beispiel, wenn Sie die Nachricht „Ich habe niedrige Produktionskosten.“ gewählt haben, bekommt Ihr zugeordneter Käufer folgende Bildschirmanzeige:



Ihr Käufer bekommt eine analoge Bildschirmanzeige, wenn Sie die Nachricht „Ich habe hohe Produktionskosten.“ gewählt haben.

Sie können nur zu Beginn der Studie dem Käufer eine Nachricht über Ihre Produktionskosten schicken. Danach haben Sie hierzu keine Möglichkeit mehr. Der Käufer wird in jeder Periode, während er den Preis und die gewünschte Qualität festlegt, wieder an Ihre Nachricht erinnert.

Ablauf einer Transaktion

Nach Bestimmung der Produktionskosten führen Sie und Ihr zugeordneter Käufer in jeder der 15 Runden eine Transaktion durch. Der Ablauf jeder Runde ist wie folgt organisiert:

1.Stufe: Preisangabe und Qualitätswunsch.
2.Stufe: Festlegung der tatsächlichen Produktqualität.
3.Stufe: Bestimmung der Einkommen.

<u>1.Stufe: Preisangabe und Qualitätswunsch</u>

Zu Beginn jeder Runde bezahlt der Käufer einen Preis und gibt einen Qualitätswunsch an. **Die Käufer müssen folgende Regeln einhalten:**

- Der Preis darf nicht kleiner als 0 und nicht höher als 100 sein:

0 ≤ Preis≤ 100.

- Die gewünschte Produktqualität darf nicht kleiner als 1 und nicht höher als 10 sein:

1 ≤ gewünschte Produktqualität ≤ 10.

Die Wahl des Preises und die Angabe der gewünschten Qualität tätigt Ihr Käufer über folgende Bildschirmanzeige:

Im obigen Beispiel der Bildschirmanzeige hat der Verkäufer dem Käufer zu Beginn der Studie die Nachricht gesendet „Ich habe niedrige Produktionskosten.“

<u>2.Stufe: Festlegung der tatsächlichen Produktqualität</u>

Daraufhin erhalten Sie den bezahlten Preis und werden über den Qualitätswunsch informiert. Dann können Sie bestimmen, welche Produktqualität Sie an Ihren Käufer liefern wollen. **Die von Ihrem Käufer gewünschte Produktqualität ist für Sie als Verkäufer nicht bindend. Sie können exakt die vom Käufer gewünschte Qualität wählen, aber auch eine höhere oder tiefere Qualität.**

Die Qualität, die Sie wählen, muss ein ganzzahliger Wert zwischen 1 und 10 sein:

$1 \leq$ tatsächliche Produktqualität ≤ 10.

Die Wahl der tatsächlichen Produktqualität tätigen Sie über die folgende Bildschirmanzeige:

In obigem Beispiel der Bildschirmanzeige ist der Verkäufer den niedrigen Kosten zugeordnet und hat zu Beginn der Studie die Nachricht gesendet „Ich habe niedrige Produktionskosten.“

Um die tatsächliche Produktqualität zu wählen, die Sie liefern möchten, geben Sie den Wert für die Qualität in das Feld „Tatsächliche Qualität:“ ein. **Bitte beachten Sie hierbei stets die mit der tatsächlichen Qualität verbundenen Kosten, die von Ihrer Zuteilung zu den hohen oder niedrigen Produktionskosten abhängen.**

6

Klicken Sie auf den „weiter“-Knopf um eine Vorschau auf Ihr Einkommen und das Einkommen Ihres Käufers zu erhalten, das sich durch Ihre gewählte Qualität ergibt. Im Rahmen dieser Vorschau haben Sie die Möglichkeit durch klicken des „Qualität Ändern“-Knopfs die Wahl Ihrer tatsächlichen Qualität in dieser Runde zu korrigieren. Durch klicken des “OK“-Knopfs machen Sie Ihre eingegebene Qualität in dieser Runde definitiv.

3.Stufe: Bestimmung der Einkommen

Ihr Einkommen hängt vom bezahlten Preis, der von Ihnen gewählten Produktqualität und den Ihnen zugewiesenen Produktionskosten ab.

Wenn Sie den niedrigen Produktionskosten zugeordnet sind, dann berechnet sich Ihr Einkommen wie folgt:

Ihr Einkommen = Preis – niedrige Produktionskosten der gewählten Produktqualität

Wenn Sie den hohen Produktionskosten zugeordnet sind, dann berechnet sich Ihr Einkommen wie folgt:

Ihr Einkommen = Preis – hohe Produktionskosten der gewählten Produktqualität

Ihr Einkommen ist demzufolge umso höher, je höher der vom Käufer angegebene Preis. Ausserdem gilt, dass Ihr Einkommen umso höher ist, je tiefer die Produktionskosten.
Wie oben beschrieben gilt, dass mehr Qualität immer mehr kostet. Darüber hinaus sind Ihre Produktionskosten für jede gewählte Qualitätsstufe höher, wenn Sie den hohen Produktionskosten zugewiesen sind, als wenn Sie den niedrigen Produktionskosten zugeordnet sind. Die genauen Werte entnehmen Sie oben stehender Tabelle oder der Tabelle und Grafik auf dem Zusatzblatt.

Das Einkommen Ihres Käufers hängt davon ab, welchen Preis er bezahlt hat, und welchen Wert das Produkt für Ihn hat. Der Wert für Ihren Käufer wird durch die von Ihnen gewählte tatsächliche Produktqualität wie folgt bestimmt:

Wert für Ihren Käufer = 10* tatsächliche Produktqualität.

Eine tabellarische und grafische Beschreibung des Werts für den Käufer für jede tatsächliche Produktqualität entnehmen Sie dem Zusatzblatt.

7

Insgesamt ergibt sich das Einkommen Ihres Käufers wie folgt:

Einkommen Ihres Käufers = Wert für Ihren Käufer – Preis
= 10*tatsächliche Produktqualität – Preis

Das Einkommen Ihres Käufers ist folglich umso höher, je höher die von Ihnen gelieferte Produktqualität, da eine höhere Qualität einen höheren Wert für Ihren Käufer ergibt. Gleichzeitig ist sein Einkommen umso höher, je tiefer der Preis, den er für das Produkt bezahlt.

Die Einkommen aller Verkäufer und Käufer werden in der oben beschriebenen Weise berechnet. Jeder Verkäufer kann also das Einkommen seines Käufers berechnen. Ein Käufer kann das Einkommen seines Verkäufers bestimmen, unter der Annahme, dass der Verkäufer hohe oder niedrige Produktionskosten hat. Da der Käufer die tatsächlichen Produktionskosten jedoch nicht eindeutig kennt (er sieht nur die Nachricht des Verkäufers), kann er das tatsächliche Einkommen des Verkäufers nicht eindeutig bestimmen.

Beachten Sie, dass Käufer und Verkäufer in jeder Periode auch Verluste erzielen können. Diese müssen aus dem jeweiligen Startgeld bzw. aus in anderen Perioden erzielten Einkommen bezahlt werden. Beachten Sie, dass eigene Verluste sowie Verluste für den anderen immer durch eigene Entscheidungen mit Sicherheit ausgeschlossen werden können.

Nachstehende Beispiele sollen die Bestimmung der Einkommen illustrieren.

Beispiel 1:
Ein Käufer bezahlt einen Preis von 35 und wünscht eine Qualität von 8. Sein Verkäufer wählt eine tatsächliche Qualität von 4.
Das Einkommen des Käufers ist: 4*10-35=5.
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen: 35-2=33.
Wenn der Verkäufer hohe Kosten hat, dann ist sein Einkommen: 35-10=25.

Beispiel 2:
Ein Käufer bezahlt einen Preis von 50 und wünscht eine Qualität von 10. Sein Verkäufer wählt eine tatsächliche Qualität von 8.
Das Einkommen des Käufers ist: 8*10-50=30.
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen: 50-10=40.
Wenn der Verkäufer hohe Kosten hat, dann ist sein Einkommen: 50-30=20.

Beispiel 3:
Ein Käufer bezahlt einen Preis von 25 und wünscht eine Qualität von 4. Sein Verkäufer wählt eine tatsächliche Qualität von 7.
Das Einkommen des Käufers ist: 7*10-25=45.
Wenn der Verkäufer hohe Kosten hat, dann ist sein Einkommen: 25-8=17.
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen: 25-25=0.

Am Ende jeder Runde wird Ihnen eine Zusammenfassung der aktuellen Runde auf dem Bildschirm präsentiert. Diese beinhaltet:

8

- Welchen Preis der Käufer angegeben hat.
- Die gewünschte Qualität des Käufers.
- Die tatsächliche Qualität, die Sie gewählt haben.
- Ihr erzieltes Einkommen in dieser Runde.

Auch Ihr Käufer erhält eine Zusammenfassung. Er wird informiert über:

- Seinen angegebenen Preis.
- Seine gewünschte Qualität.
- Die tatsächliche Qualität, die Sie gewählt haben.
- Sein erzieltes Einkommen in dieser Runde.

Nachdem die Zusammenfassung zu sehen ist, ist die aktuelle Runde abgeschlossen. Danach beginnt die nächste Runde. Insgesamt besteht die Studie aus 15 Runden, und Sie sind in jeder der 15 Runden demselben Käufer zugeordnet.

Die Studie beginnt erst dann, wenn alle TeilnehmerInnen mit dem Ablauf der Studie vollständig vertraut sind. Um dies sicher zu stellen, bitten wir Sie unten stehende Übungsaufgaben zu lösen.

9

Übungsaufgaben

Bitte lösen Sie diese **Aufgaben vollständig und unter Angabe des Lösungswegs**. Wenn Sie Fragen haben, wenden Sie sich bitte an die Leiter des Experiments.

Aufgabe 1:
Der bezahlte Preis des Käufers ist 55 und seine gewünschte Qualität ist 9. Der zugewiesene Verkäufer wählt eine tatsächliche Qualität von 8.

Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Aufgabe 2:
Der bezahlte Preis des Käufers ist 60 und seine gewünschte Qualität ist 9. Der zugewiesene Verkäufer wählt eine tatsächliche Qualität von 5.

Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Aufgabe 3:
Der bezahlte Preis des Käufers ist 9 und seine gewünschte Qualität ist 4. Der zugewiesene Verkäufer wählt eine tatsächliche Qualität von 1.

Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Aufgabe 4:
Der bezahlte Preis des Käufers ist 40 und seine gewünschte Qualität ist 10. Der zugewiesene Verkäufer wählt eine tatsächliche Qualität von 8.

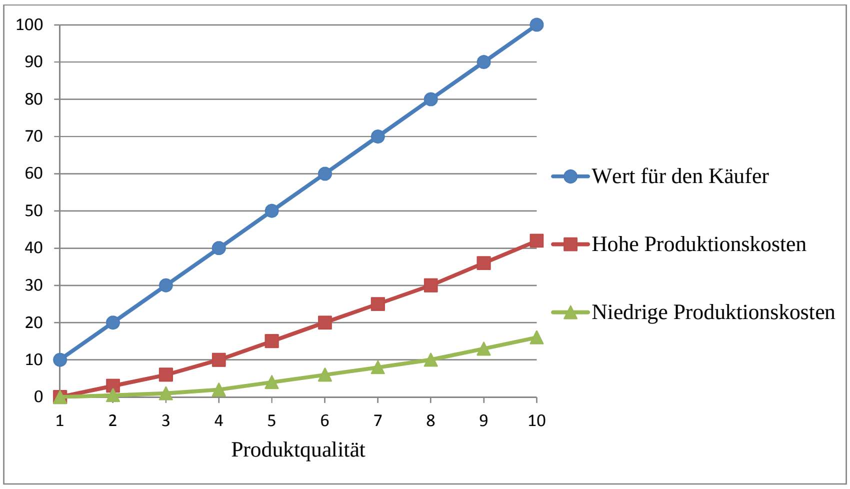
Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Wenn Sie mit dem Lösen der Übungsaufgaben fertig sind, empfehlen wir Ihnen, sich die Aufgaben und deren Lösungen noch einmal anzusehen. Anschliessend überlegen Sie sich bitte, welche Entscheidungen Sie im Experiment treffen wollen.

10

Zusatzblatt

Produktqualität	1	2	3	4	5	6	7	8	9	10
Niedrige Produktionskosten	0	0.5	1	2	4	6	8	10	13	16
Hohe Produktionskosten	0	3	6	10	15	20	25	30	36	42
Wert für den Käufer	10	20	30	40	50	60	70	80	90	100



11

B Appendix to Chapter 3

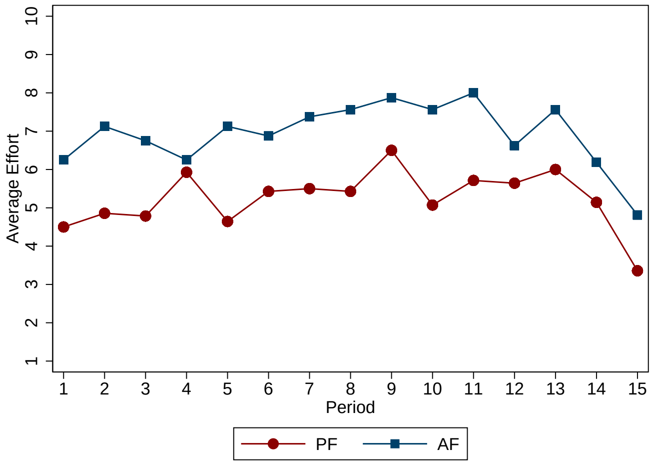
B.1 Regressions and Figures

Table B.1: Wage-Effort Regressions

	(1)	(2)
Constant	-11.41^{***}	-4.70^{***}
	(2.35)	(1.26)
Effort	7.96^{***}	7.21^{***}
	(0.31)	(0.19)
$\mathbf{1}_{[\mathrm{P}]}$	8.76^{***}	16.25^{***}
	(2.00)	(1.66)
N	30	128
Adj. R^2	0.94	0.87
Treatment order	F	X

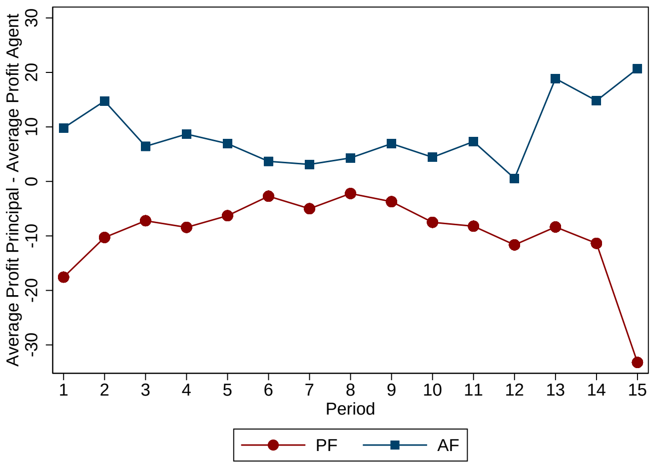
Notes: Column (1) relates to data observed in the fixed order treatments PF and AF and column (2) to data elicited in the alternating order conditions PX and AX. The dependent variable in all of the above regressions is wage. The analysis is performed on average wages and average efforts as described below Figure 3.2 and Figure 3.4. Average efforts were subtracted by one such that they can take values in $[0, 9]$. $\mathbf{1}_{[P]}$ represents a dummy which equals one if the transactions were conducted under the order P and zero otherwise: in column (1), the dummy equals one if the observation was made in the context of treatment PF and zero if it was observed in AF; in column (2), the dummy equals one if the vector of average effort and price was computed over transactions conducted in the order P and zero if it was computed over rounds which took place in the order A. Standard errors are reported in parentheses and are robust in column (1) and clustered on the pair level in column (2). *** denotes significance at the 1 percent level.

Figure B.1: Effort in the Fixed Order Treatments



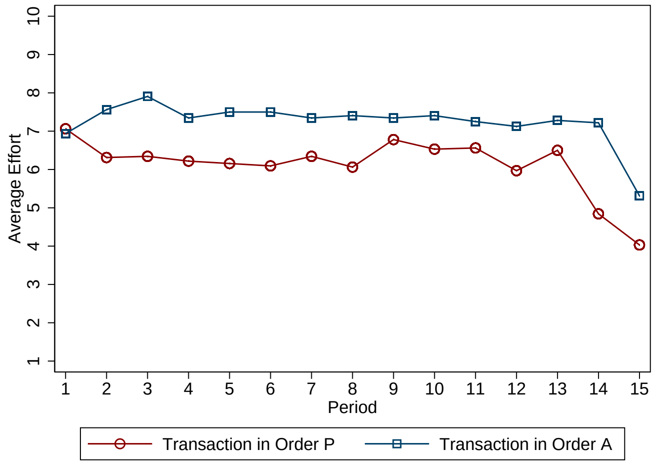
Note: One data point marked by a dot (square) indicates the average effort taken over all agents within treatment PF (AF) in a given period.

Figure B.2: Difference in Profits in the Fixed Order Treatments



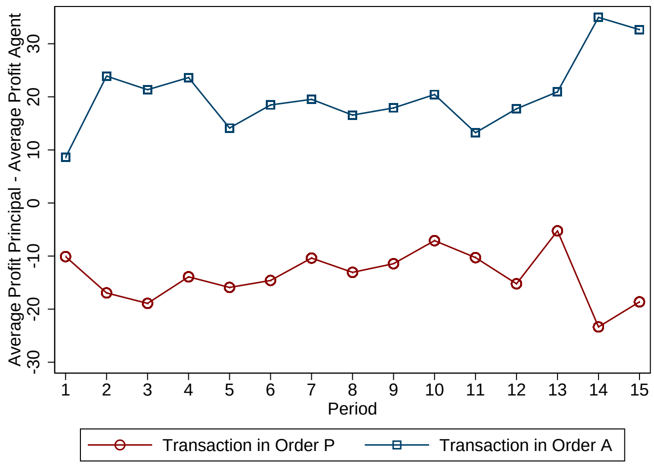
Note: One data point marked by a dot (square) represents the difference between the average of principals' profits per transaction and the average of agents' payoffs per round where the averages are computed for a given period and over all principals or agents, respectively, observed in treatment PF (AF).

Figure B.3: Effort in the Alternating Order Treatments



Note: One data point marked by a circle (hollow square) represents the average effort computed over all agents in the treatments PX and AX for whom the transaction in a given period took place in the order P(A); e.g. in period 3, all agents in PX (AX) are concerned.

Figure B.4: Difference in Profits in the Alternating Order Treatments



Note: One data point marked with a circle (hollow square) represents the difference between the principals' average profits and the agents' average profits where the averages are computed over all principals and agents, respectively, in the alternating order treatments, for whom their transaction was conducted under the order P(A) at the given period; e.g. in period 5, the data point corresponds to data elicited under condition PX(AX).

B.2 Instructions

B.2.1 Instructions for Sellers in Treatment AF

Allgemeine Erklärungen für Verkäufer

Sie nehmen nun an einer wirtschaftswissenschaftlichen Studie teil, die von diversen Forschungsförderungsstellen finanziert wird. Sie können dabei - abhängig von Ihren Entscheidungen – Geld verdienen. Es ist daher sehr wichtig, dass Sie diese Erklärungen genau durchlesen.

Diese Instruktionen dienen ausschliesslich Ihrer privaten Information. **Während der Studie herrscht ein absolutes Kommunikationsverbot.** Wenn Sie Fragen haben, dann richten Sie diese bitte an uns. Die Nichtbeachtung dieser Regel führt zum Ausschluss aus der Studie und von allen Zahlungen.

Zu Beginn der Studie erhalten Sie ein Startgeld von 10 Franken. Während der Studie sprechen wir nicht von Franken, sondern von Punkten. Im Verlauf der Studie können Sie einen weiteren Geldbetrag verdienen, indem Sie Punkte erzielen. Ihr gesamtes Einkommen wird also zunächst in Punkten berechnet. Die von Ihnen während der Studie erzielte Gesamtpunktezahl wird dann am Ende in Franken umgerechnet, dabei gilt

10 Punkte = 1 Franken.

Sollten Sie im Laufe der Studie Verluste machen, werden allfällige Verluste mit Ihrem Startgeld verrechnet. **Sie können jedoch Verluste immer durch eigene Entscheidungen <u>mit Sicherheit</u> ausschliessen!** Am Ende bekommen Sie von uns die während der Studie verdiente Punktezahl plus die 10 Franken Startgeld in **bar** ausbezahlt.

Zu Beginn der Studie wurden die TeilnehmerInnen zufällig in zwei Gruppen aufgeteilt: Käufer und Verkäufer. **Sie sind während der gesamten Studie ein Verkäufer.** Jeder Verkäufer wurde zufällig einem Käufer zugeordnet. Die Studie besteht aus 15 Runden. Sie sind in jeder der 15 Runden dem gleichen Käufer zugeordnet. Keiner der anderen Studienteilnehmer wird Ihre genaue Zuordnung oder Ihre Entscheidungen erfahren. Die Anonymität bleibt also gewahrt.

1

Kurzübersicht über die Studie

Zu Beginn der Studie wurde jedem Verkäufer zufällig ein Käufer zugeordnet. Die Studie umfasst 15 Runden, in denen die Zuordnung von Verkäufer und Käufer fixiert bleibt. In jeder Runde führt jeder Verkäufer mit seinem zugeordneten Käufer eine Transaktion durch: Der Verkäufer bestimmt die Qualität eines Produkts für das der Käufer einen Preis bezahlt. Der Käufer erzielt durch die Transaktion einen Gewinn, wenn er für das Produkt weniger bezahlt, als das Produkt ihm wert ist. Wie hoch der Wert des Produktes für den Käufer ist, hängt von der Qualität des Produktes ab. Der Verkäufer erzielt durch die Transaktion einen Gewinn, wenn er einen Preis erhält, der seine Produktionskosten übersteigt. Produktionskosten entstehen durch die Bereitstellung von Produktqualität. Eine höhere Qualität ist immer mit höheren Produktionskosten verbunden.

<u>Bestimmung der Produktionskosten</u>

Bevor Käufer und Verkäufer Transaktionen durchführen, werden die genauen Produktionskosten bestimmt. Ein Verkäufer kann entweder **hohe oder niedrige Produktionskosten** haben. Unter den hohen Produktionskosten kostet jede Qualitätsstufe mehr als unter den niedrigen Produktionskosten. Die eine Hälfte aller Verkäufer wird zufällig den hohen Produktionskosten zugewiesen, die andere Hälfte aller Verkäufer den niedrigen Produktionskosten zugeordnet. Diese Zuweisung bleibt über die gesamte Studie hinweg bestehen.

Sie werden zu Beginn der Studie darüber informiert, ob Sie den hohen oder niedrigen Produktionskosten zugeordnet sind. Der Ihnen zugeordnete Käufer wird ebenfalls über Ihre genauen Produktionskosten informiert.

<u>Ablauf einer Transaktion</u>

Nach Bestimmung der Produktionskosten führen Sie und Ihr zugeordneter Käufer in jeder der 15 Runden eine Transaktion durch. Der Ablauf ist folgendermassen organisiert:

1.Stufe: Sie wählen die Produktqualität, die Sie an Ihren Käufer liefern wollen, und geben einen gewünschten Preis an.

2.Stufe: Ihr Käufer erhält daraufhin die von Ihnen gewählte Produktqualität, wird über Ihren gewünschten Preis informiert und wählt den tatsächlichen Preis, den er Ihnen bezahlen möchte. **Der von Ihnen gewünschte Preis ist für die Wahl des tatsächlichen Preises durch Ihren Käufer nicht bindend.**

3.Stufe: Sie erhalten den von Ihrem Käufer gewählten Preis und werden über Ihr Einkommen in der aktuellen Runde informiert. Danach beginnt die nächste Runde.

Die Einkommen aus allen 15 Runden werden am Ende der Studie zusammengezählt, in Franken umgerechnet und zusammen mit dem Startgeld bar ausbezahlt.
Auf den nächsten Seiten beschreiben wir den Ablauf der Bestimmung der Produktionskosten und die Transaktion weiter im Detail.

2

Information über den genauen Ablauf der Studie

Bestimmung der Produktionskosten

Die Produktionskosten des Verkäufers hängen von der Produktqualität ab. Die Produktqualität kann nicht kleiner als 1 und nicht höher als 10 sein:

1 ≤ Produktqualität ≤ 10.

Die Produktionskosten sind umso höher, je höher die gewählte Qualität. **Die Produktionskosten hängen zudem davon ab, ob der Verkäufer den hohen oder niedrigen Produktionskosten zugewiesen ist.** Wenn Sie den hohen Produktionskosten zugewiesen sind, dann sind Ihre mit jeder Qualitätsstufe verbundenen Produktionskosten höher als wenn Sie den niedrigen Produktionskosten zugeordnet sind. Die niedrigen und hohen Produktionskosten sind für jede Produktqualität in der nachstehenden Tabelle beschrieben. Die Tabelle sowie eine grafische Beschreibung finden Sie auch auf dem Zusatzblatt.

Produktqualität	1	2	3	4	5	6	7	8	9	10
Niedrige Produktionskosten	0	0.5	1	2	4	6	8	10	13	16
Hohe Produktionskosten	0	3	6	10	15	20	25	30	36	42

Der Computer weist zu Beginn der Studie per Zufall die eine Hälfte aller Verkäufer den hohen Produktionskosten und die andere Hälfte den niedrigen Produktionskosten zu.
Sie werden auf dem Bildschirm über Ihre tatsächlichen Produktionskosten informiert. Die tatsächlichen Produktionskosten bleiben nach der einmaligen Zuordnung zu Beginn der ersten Periode in jeder der 15 Runden gleich. Ihr zugeordneter Käufer wird ebenfalls über Ihre tatsächlich zugewiesenen Produktionskosten informiert.
Zum Beispiel, wenn Sie den niedrigen Kosten zugeordnet sind, sehen sie folgende Information auf Ihrem Bildschirm:

> Der Computer hat per Zufall die eine Hälfte aller Verkäufer den hohen Produktionskosten und die andere Hälfte der Verkäufer den niedrigen Produktionskosten zugewiesen.
>
> Sie sind in dieser Studie den **niedrigen** Produktionskosten zugeordnet.

Im Fall der Zuordnung zu den niedrigen Kosten bekommt Ihr zugeordneter Käufer folgende Information:

3

Der Computer hat per Zufall die eine Hälfte aller Verkäufer den hohen Produktionskosten und die andere Hälfte der Verkäufer den niedrigen Produktionskosten zugewiesen.

Ihr Verkäufer hat **niedrige** Produktionskosten.

Wenn Sie den hohen Kosten zugeordnet sind, sehen Sie und Ihr Käufer analoge Information auf Ihren Bildschirmen.

Sie und Ihr Käufer werden in jeder Runde, während sie Ihre Entscheidungen treffen, an die Ihnen zugewiesenen Produktionskosten erinnert.

Ablauf einer Transaktion

Nach Bestimmung der Produktionskosten führen Sie und Ihr zugeordneter Käufer in jeder der 15 Runden eine Transaktion durch. Der Ablauf jeder Runde ist wie folgt organisiert:

1.Stufe: Festlegung der Qualität und Angabe eines gewünschten Preises.
2.Stufe: Festlegung des tatsächlichen Preises.
3.Stufe: Bestimmung der Einkommen.

1.**Stufe: Festlegung der Qualität und Angabe eines gewünschten Preises.**

Zu Beginn jeder Runde legen Sie eine Produktqualität fest und geben einen gewünschten Preis an. **Sie müssen müssen folgende Regeln einhalten:**

- Die Qualität, die Sie wählen, muss ein ganzzahliger Wert zwischen 1 und 10 sein:

$$\mathbf{1 \leq \text{Qualität} \leq 10.}$$

- Ihr gewünschter Preis darf nicht kleiner als 0 und nicht höher als 100 sein:

$$\mathbf{0 \leq \text{gewünschter Preis} \leq 100.}$$

Die Wahl der Qualität und die Angabe des gewünschten Preises tätigen Sie über folgende Bildschirmanzeige:

4

Im obigen Beispiel der Bildschirmanzeige ist der Verkäufer den niedrigen Kosten zugeordnet.

Geben Sie die Produktqualität, die Sie liefern möchten, und den gewünschten Preis in die jeweils dafür vorgesehenen Felder ein. **Beachten Sie stets die mit der Qualität verbundenen Kosten, die von Ihrer Zuteilung zu den hohen oder niedrigen Produktionskosten abhängen.** Machen Sie Ihre Eingaben durch klicken des „OK"-Knopfs in dieser Runde definitiv.

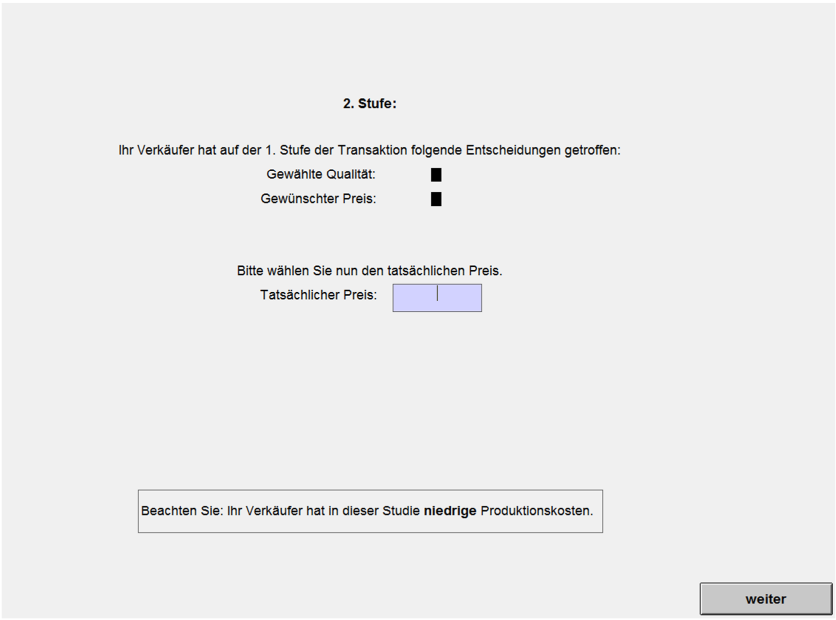
2.Stufe: Festlegung des tatsächlichen Preises

Daraufhin erhält Ihr Käufer die bereit gestellte Qualität und wird über Ihren gewünschten Preis informiert. Dann kann Ihr Käufer bestimmen, welchen Preis er tatsächlich bezahlen will. **Der von Ihnen gewünschte Preis ist für Ihren Käufer nicht bindend. Ihr Käufer kann exakt den von Ihnen gewünschten Preis wählen, aber auch einen höheren oder niedrigeren Preis.**

Der Preis, den Ihr Käufer wählt, darf nicht kleiner als 0 und nicht höher als 100 sein:

$$\mathbf{0 \leq \text{tatsächlicher Preis} \leq 100.}$$

Die Wahl des tatsächlichen Preises tätigt Ihr Käufer über die folgende Bildschirmanzeige:



In obigem Beispiel der Bildschirmanzeige ist der Verkäufer den niedrigen Kosten zugeordnet.

3.Stufe: Bestimmung der Einkommen

Ihr Einkommen hängt vom Preis, den Ihr Käufer bezahlt hat, der von Ihnen gewählten Produktqualität und den Ihnen zugewiesenen Produktionskosten ab.

Wenn Sie den niedrigen Produktionskosten zugeordnet sind, dann berechnet sich Ihr Einkommen wie folgt:

Ihr Einkommen, wenn Sie niedrige Kosten haben
= tatsächlicher Preis – niedrige Produktionskosten der gewählten Qualität

Wenn Sie den hohen Produktionskosten zugeordnet sind, dann berechnet sich Ihr Einkommen wie folgt:

Ihr Einkommen, wenn Sie hohe Kosten haben
= tatsächlicher Preis – hohe Produktionskosten der gewählten Qualität

6

Ihr Einkommen ist demzufolge umso höher, je höher der von Ihrem Käufer gewählte tatsächliche Preis. Ausserdem gilt, dass Ihr Einkommen umso höher ist, je tiefer die Produktionskosten.
Wie oben beschrieben gilt, dass mehr Qualität immer mehr kostet. Darüber hinaus sind Ihre Produktionskosten für jede gewählte Qualitätsstufe höher, wenn Sie den hohen Produktionskosten zugewiesen sind, als wenn Sie den niedrigen Produktionskosten zugeordnet sind. Die genauen Werte entnehmen Sie oben stehender Tabelle oder der Tabelle und Grafik auf dem Zusatzblatt.

Das Einkommen Ihres Käufers hängt davon ab, welchen Preis er bezahlt hat, und welchen Wert das Produkt für Ihn hat. Der Wert für Ihren Käufer wird durch die von Ihnen gewählte Produktqualität wie folgt bestimmt:

Wert für Ihren Käufer = 10* gewählte Qualität.

Eine tabellarische und grafische Beschreibung des Werts für den Käufer für jede gewählte Produktqualität entnehmen Sie dem Zusatzblatt.

Insgesamt ergibt sich das Einkommen Ihres Käufers wie folgt:

Einkommen Ihres Käufers = Wert für Ihren Käufer – tatsächlicher Preis
= 10* gewählte Qualität – tatsächlicher Preis

Das Einkommen Ihres Käufers ist folglich umso höher, je höher die von Ihnen gelieferte Produktqualität, da eine höhere Qualität einen höheren Wert für Ihren Käufer ergibt. Gleichzeitig ist sein Einkommen umso höher, je tiefer der Preis, den er für das Produkt bezahlt.

Die Einkommen aller Verkäufer und Käufer werden in der oben beschriebenen Weise berechnet. Jeder Verkäufer kann also das Einkommen seines Käufers berechnen. Zudem kann jeder Käufer das Einkommen seines Verkäufers bestimmen, unter Berücksichtigung der Zuordnung seines Verkäufers zu den hohen oder niedrigen Produktionskosten.

Beachten Sie, dass Käufer und Verkäufer in jeder Periode auch Verluste erzielen können. Diese müssen aus dem jeweiligen Startgeld bzw. aus in anderen Perioden erzielten Einkommen bezahlt werden. Beachten Sie, dass eigene Verluste sowie Verluste für den anderen immer durch eigene Entscheidungen mit Sicherheit ausgeschlossen werden können.

Nachstehende Beispiele sollen die Bestimmung der Einkommen illustrieren.

Beispiel 1:
Ein Verkäufer wählt eine Qualität von 8 und wünscht einen Preis von 60. Sein Käufer bezahlt einen Preis von 50.
Das Einkommen des Käufers ist: 8*10-50=30.
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen: 50-10=40.
Wenn der Verkäufer hohe Kosten hat, dann ist sein Einkommen: 50-30=20.

7

Beispiel 2:
Ein Verkäufer wählt eine Qualität von 4 und wünscht einen Preis von 20. Sein Käufer bezahlt einen Preis von 30.
Das Einkommen des Käufers ist: 4*10-30=10.
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen: 30-2=28.
Wenn der Verkäufer hohe Kosten hat, dann ist sein Einkommen: 30-10=20.

Beispiel 3:
Ein Verkäufer wählt eine Qualität von 7 und wünscht einen Preis von 40. Sein Käufer bezahlt einen Preis von 30.
Das Einkommen des Käufers ist: 7*10-30=40.
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen: 30-8=22.
Wenn der Verkäufer hohe Kosten hat, dann ist sein Einkommen: 30-25=5.

Am Ende jeder Runde wird Ihnen und Ihrem Käufer eine Zusammenfassung der aktuellen Runde auf dem Bildschirm präsentiert. Diese beinhaltet:
- Die gewählte Qualität.
- Den gewünschten Preis.
- Den tatsächlich bezahlten Preis.

Darüber hinaus werden Sie über Ihr Einkommen und Ihr Käufer ebenfalls über sein eigenes Einkommen in dieser Runde informiert.

Nachdem die Zusammenfassung zu sehen ist, ist die aktuelle Runde abgeschlossen. Danach beginnt die nächste Runde. Insgesamt besteht die Studie aus 15 Runden, und Sie sind in jeder der 15 Runden demselben Käufer zugeordnet.

Die Studie beginnt erst dann, wenn alle TeilnehmerInnen mit dem Ablauf der Studie vollständig vertraut sind. Um dies sicher zu stellen, bitten wir Sie die auf der nächsten Seite stehenden Übungsaufgaben zu lösen.

Übungsaufgaben

Bitte lösen Sie diese **Aufgaben vollständig und unter Angabe des Lösungswegs**. Wenn Sie Fragen haben, wenden Sie sich bitte an die Leiter der Studie.

Aufgabe 1:
Ein Verkäufer wählt eine Qualität von 10 und wünscht einen Preis von 70. Sein Käufer bezahlt einen Preis von 60.

Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Aufgabe 2:
Ein Verkäufer wählt eine Qualität von 1 und wünscht einen Preis von 4. Sein Käufer bezahlt einen Preis von 5.

Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Aufgabe 3:
Ein Verkäufer wählt eine Qualität von 5 und wünscht einen Preis von 35. Sein Käufer bezahlt einen Preis von 35.

Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Aufgabe 4:
Ein Verkäufer wählt eine Qualität von 7 und wünscht einen Preis von 60. Sein Käufer bezahlt einen Preis von 35.

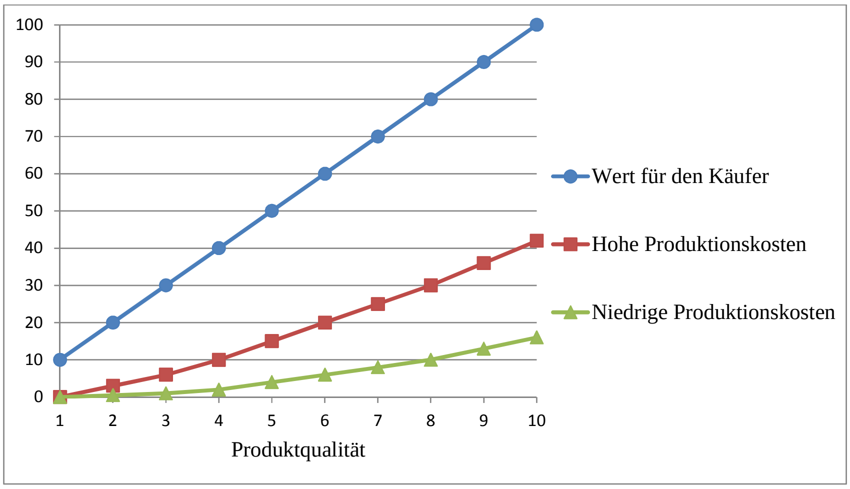
Das Einkommen des Käufers ist:
Wenn der Verkäufer niedrige Kosten hat, dann ist sein Einkommen:
Wenn der Verkäufer hohe Kosten zugeordnet hat, dann ist sein Einkommen:

Wenn Sie mit dem Lösen der Übungsaufgaben fertig sind, empfehlen wir Ihnen, sich die Aufgaben und deren Lösungen noch einmal anzusehen. Anschliessend überlegen Sie sich bitte, welche Entscheidungen Sie in der Studie treffen wollen.

9

Zusatzblatt

Produktqualität	1	2	3	4	5	6	7	8	9	10
Niedrige Produktionskosten	0	0.5	1	2	4	6	8	10	13	16
Hohe Produktionskosten	0	3	6	10	15	20	25	30	36	42
Wert für den Käufer	10	20	30	40	50	60	70	80	90	100



10

B.2.2 Instructions for Buyers in Treatment AX

Allgemeine Erklärungen für Käufer

Sie nehmen nun an einer wirtschaftswissenschaftlichen Studie teil, die von diversen Forschungsförderungsstellen finanziert wird. Sie können dabei - abhängig von Ihren Entscheidungen – Geld verdienen. Es ist daher sehr wichtig, dass Sie diese Erklärungen genau durchlesen.

Diese Instruktionen dienen ausschliesslich Ihrer privaten Information. Während der Studie herrscht ein absolutes Kommunikationsverbot. Wenn Sie Fragen haben, dann richten Sie diese bitte an uns. Die Nichtbeachtung dieser Regel führt zum Ausschluss aus der Studie und von allen Zahlungen.

Zu Beginn der Studie wurden die TeilnehmerInnen zufällig in zwei Gruppen aufgeteilt: Käufer und Verkäufer. **Sie sind während der gesamten Studie ein Käufer.** Sie werden im Rahmen dieser Studie an vier Interaktionen mit vier verschiedenen Verkäufern teilnehmen. In jeder Interaktion sind Sie jeweils einem Verkäufer fix zugeordnet.

Zu Beginn der Studie erhalten Sie ein Startgeld von 10 Franken. Im Rahmen jeder Interaktion können Sie zusätzlich Punkte erzielen. Durch Erwerb von Punkten können Sie weiteres Geld verdienen. Dabei wird genau eine Ihrer vier Interaktionen zufällig ausgewählt. Alle Punkte, die Sie im Rahmen dieser zufällig ausgewählten Interaktion erzielt haben, werden wie folgt in Franken umgerechnet:

10 Punkte = 1 Franken.

Sollten Verluste entstehen, werden diese mit Ihrem Startgeld verrechnet. Sie können jedoch Verluste immer durch eigene Entscheidungen mit Sicherheit ausschliessen. Am Ende der Studie bekommen Sie alle Punkte, die Sie im Rahmen der zufällig ausgewählten Interaktion erzielt haben, plus die 10 Franken Startgeld in **bar** ausbezahlt.

Keiner der anderen Studienteilnehmer wird die genaue Zuordnung und die Entscheidungen von Ihnen und Ihrem, in einer Interaktion zugeteilten, Verkäufer erfahren. Die Anonymität bleibt also gewahrt.

1

Kurzübersicht über die Studie

Interaktion

Die Studie besteht aus vier Interaktionen. Zu Beginn einer Interaktion wird jedem Verkäufer zufällig ein Käufer zugeordnet. Dabei wird ausgeschlossen, dass die beiden Parteien bereits in einer vergangenen Interaktion zugeordnet waren. **Eine Interaktion umfasst 15 Runden**, in denen die Zuordnung von Käufer und Verkäufer fixiert bleibt. **In jeder Runde einer Interaktion führen ein Käufer und sein zugeordneter Verkäufer eine Transaktion durch.**

Ablauf einer Transaktion

In jeder Runde führen die zugeordneten Parteien eine Transaktion durch: Der Verkäufer bestimmt die Qualität eines Produkts, für das der Käufer einen Preis bezahlt. Der Käufer erzielt durch die Transaktion einen Gewinn, wenn er für das Produkt weniger bezahlt, als das Produkt ihm wert ist. Wie hoch der Wert des Produktes für den Käufer ist, hängt von der Qualität des Produktes ab. Der Verkäufer erzielt durch die Transaktion einen Gewinn, wenn er einen Preis erhält, der seine Produktionskosten übersteigt. Produktionskosten entstehen durch die Bereitstellung von Produktqualität. Eine höhere Qualität ist immer mit höheren Produktionskosten verbunden.

In jeder ungeraden Runde (1,3,5,…,15) ist eine Transaktion in folgende Stufen eingeteilt:

1.Stufe: Der Verkäufer wählt eine Produktqualität und gibt einen gewünschten Preis an.

2.Stufe: Der Käufer erhält daraufhin die gewählte Produktqualität, wird über den gewünschten Preis informiert und wählt den tatsächlichen Preis, den er bezahlen möchte. **Der vom Verkäufer gewünschte Preis ist für die Wahl des tatsächlichen Preises durch den Käufer nicht bindend.**

3.Stufe: Der Verkäufer erhält den von seinem Käufer gewählten Preis. Käufer und Verkäufer werden jeweils über die erzielten Punkte in der aktuellen Runde informiert. Danach beginnt die nächste Runde.

In jeder geraden Runde (2,4,6,…,14) ist eine Transaktion in folgende Stufen eingeteilt:

1.Stufe: Der Käufer bezahlt einen Preis und gibt eine gewünschte Produktqualität an.

2.Stufe: Der Verkäufer erhält daraufhin den bezahlten Preis, wird über die gewünschte Qualität informiert, und wählt die tatsächliche Produktqualität, die er liefern möchte. **Die gewünschte Qualität des Käufers ist für die Wahl der tatsächlichen Qualität durch den Verkäufer nicht bindend.**

3.Stufe: Der Käufer erhält die von seinem Verkäufer gewählte Qualität. Käufer und Verkäufer werden jeweils über die erzielten Punkte in der aktuellen Runde informiert. Danach beginnt die nächste Runde.

Nachstehend finden Sie weitere wichtige Details zu einer Transaktion.

2

Details zu einer Transaktion

In jeder Runde müssen Käufer und Verkäufer bei ihren Entscheidungen folgende Regeln einhalten:

- Die tatsächliche Qualität, die ein Verkäufer wählt, muss ein ganzzahliger Wert zwischen 1 und 10 sein: **1 ≤ tatsächliche Qualität ≤ 10.**

- Der tatsächliche Preis, den ein Käufer wählt, darf nicht kleiner als 0 und nicht höher als 100 sein: **0 ≤ tatsächlicher Preis ≤ 100.**

Auf der 1. Stufe jeder geraden Runde muss ein Käufer zusätzlich berücksichtigen:

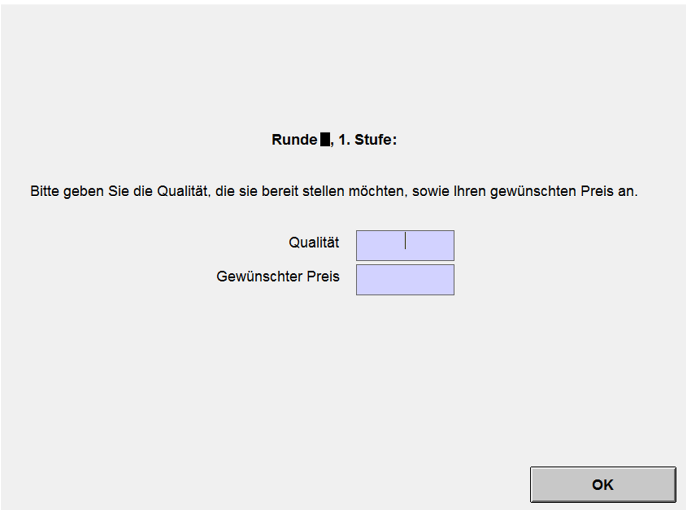
- Die gewünschte Qualität muss ein ganzzahliger Wert zwischen 1 und 10 sein: **1 ≤ gewünschte Qualität ≤ 10.**

Auf der 1. Stufe jeder ungeraden Runde muss der Verkäufer zusätzlich berücksichtigen:

- Der gewünschte Preis darf nicht kleiner als 0 und nicht höher als 100 sein: **0 ≤ gewünschter Preis ≤ 100.**

<u>1.Stufe einer Transaktion</u>

In jeder ungeraden Runde gibt der Verkäufer eine Qualität und einen gewünschten Preis über nachstehend abgebildete Bildschirmanzeige an:



Durch klicken des „OK“-Knopfes macht ein Verkäufer seine Eingaben in der Runde definitiv.

3

In jeder geraden Runde wählt der Käufer einen Preis und gibt eine gewünschte Qualität an. Die Eingabe dieser Entscheidungen findet über eine analoge Bildschirmanzeige statt und wird ebenfalls durch klicken eines „OK“-Knopfes definitiv gemacht.

2.Stufe der Transaktion

In jeder ungeraden Runde erhält der Käufer daraufhin die vom Verkäufer gewählte Qualität und wird über dessen gewünschten Preis informiert. Der Käufer kann dann den Preis bestimmen, den er tatsächlich bezahlen will. **Der vom Verkäufer gewünschte Preis ist für den Käufer dabei nicht bindend. Ein Käufer kann exakt den vom Verkäufer gewünschten Preis wählen, aber auch einen höheren oder niedrigeren Preis.**
Die Wahl des tatsächlichen Preises tätigt ein Käufer über die folgende Bildschirmanzeige:

Durch klicken des „weiter“- Knopfes erhält ein Käufer eine Vorschau auf die erzielten Einkommen in dieser Runde. Ein Käufer hat dann die Möglichkeit die Wahl des tatsächlichen Preises noch einmal zu korrigieren, bevor er seine Entscheidung definitiv macht.

In jeder geraden Runde erhält der Verkäufer den vom Käufer bezahlten Preis und wird über dessen gewünschte Qualität informiert. Der Verkäufer kann dann die Qualität bestimmen, die er tatsächlich liefern will. **Die vom Käufer gewünschte Qualität ist für den Verkäufer dabei nicht bindend. Ein Verkäufer kann exakt die vom Käufer gewünschte Qualität wählen, aber auch eine höhere oder niedrigere Qualität.**
Die Eingabe der tatsächlichen Qualität tätigt der Verkäufer über eine Bildschirmanzeige, die analog zur oben Abgebildeten ist. Ein Verkäufer erhält danach ebenso eine Vorschau auf die erzielten Einkommen in dieser Runde und hat die Möglichkeit seine Eingabe noch einmal zu korrigieren, bevor er seine Entscheidung in der aktuellen Runde definitiv macht.

4

<u>3.Stufe der Transaktion</u>

Das Einkommen eines Käufers hängt ab vom tatsächlich bezahlten Preis und dem Wert, den das Produkt für ihn hat. Der Wert des Produkts für einen Käufer wird durch die von seinem Verkäufer gewählte tatsächliche Produktqualität wie folgt bestimmt:

Wert für den Käufer = 10* tatsächliche Qualität.

Insgesamt ergibt sich das Einkommen eines Käufers wie folgt:

Einkommen eines Käufers = Wert für den Käufer – tatsächlicher Preis
= 10* tatsächliche Qualität – tatsächlicher Preis

Das Einkommen eines Käufers ist folglich umso höher, je höher die von seinem Verkäufer gelieferte Produktqualität, da eine höhere Qualität einen höheren Wert für den Käufer ergibt. Gleichzeitig ist das Einkommen eines Käufers umso höher, je tiefer der Preis, den er für das Produkt bezahlt.

Das Einkommen eines Verkäufers hängt ab vom tatsächlich bezahlten Preis und den Produktionskosten, die durch die Wahl der tatsächlichen Produktqualität entstehen:

Einkommen eines Verkäufers =
tatsächlicher Preis – Produktionskosten der tatsächlichen Qualität

Das Einkommen eines Verkäufers ist demzufolge umso höher, je höher der von seinem Käufer gewählte tatsächliche Preis. Ausserdem gilt, dass das Einkommen eines Verkäufers umso höher ist, je niedriger die anfallenden Produktionskosten. Wie Sie nachstehender Tabelle entnehmen können, ist eine höhere Qualität immer mit höheren Produktionskosten verbunden.

Qualität	1	2	3	4	5	6	7	8	9	10
Produktionskosten	0	3	6	10	15	20	25	30	36	42

Dem beigelegten Zusatzblatt können Sie eine grafische Darstellung der Produktionskosten sowie des Werts für den Käufer entnehmen.
Die Einkommen aller Verkäufer und Käufer werden in der oben beschriebenen Weise berechnet. Jeder Verkäufer kann also das Einkommen seines Käufers berechnen. Zudem kann jeder Käufer das Einkommen seines Verkäufers bestimmen.

Beachten Sie, dass Käufer und Verkäufer in jeder Runde auch Verluste erzielen können. Beachten Sie, dass eigene Verluste sowie Verluste für den anderen immer durch eigene Entscheidungen mit Sicherheit ausgeschlossen werden können.

Die erzielten Einkommen werden am Ende der Studie wie folgt in Franken umgerechnet: Genau eine Ihrer vier Interaktionen wird zufällig ausgewählt. Alle Einkommen, die Sie in den 15 Runden dieser zufällig ausgewählten Interaktion erzielt haben, werden am Ende der Studie zusammen gezählt, in Franken umgerechnet und zusammen mit dem Startgeld in bar ausbezahlt.

Nachstehende Beispiele sollen die Bestimmung der Einkommen illustrieren.

Beispiel 1:
In Runde 2 bezahlt ein Käufer einen Preis von 50 und wünscht eine Qualität von 10. Sein Verkäufer wählt eine Qualität von 8.
Das Einkommen des Käufers ist: 8*10-50=30.
Das Einkommen des Verkäufers ist: 50-30=20.

Beispiel 2:
In Runde 7 wählt ein Verkäufer eine Qualität von 6 und wünscht einen Preis von 50. Sein Käufer bezahlt einen Preis von 30.
Das Einkommen des Käufers ist: 6*10-30=30.
Das Einkommen des Verkäufers ist: 30-20=10.

Beispiel 3:
In Runde 11 wählt ein Verkäufer eine Qualität von 7 und wünscht einen Preis von 50. Sein Käufer bezahlt einen Preis von 40.
Das Einkommen des Käufers ist: 7*10-40=30.
Das Einkommen des Verkäufers ist: 40-25=15.

Am Ende jeder Runde wird jedem Käufer und Verkäufer eine Zusammenfassung der Entscheidungen beider Parteien sowie das jeweilige Einkommen in der aktuellen Runde auf dem Bildschirm präsentiert. Klicken Sie auf den „weiter zur nächsten Runde"-Knopf um die nächste Runde zu beginnen.

Die Studie beginnt erst dann, wenn alle TeilnehmerInnen mit dem Ablauf der Studie vollständig vertraut sind. Um dies sicher zu stellen, bitten wir Sie die auf der nächsten Seite stehenden Übungsaufgaben zu lösen.

6

Übungsaufgaben

Bitte lösen Sie diese **Aufgaben vollständig und unter Angabe des Lösungswegs**. Wenn Sie Fragen haben, wenden Sie sich bitte an die Leiter der Studie.

Aufgabe 1:
In Runde 2 wählt ein Käufer einen Preis 45 und wünscht eine Qualität von 10. Der Verkäufer wählt eine Qualität von 8.

Das Einkommen des Käufers ist:
Das Einkommen des Verkäufers ist:

Aufgabe 2:
In Runde 5 wählt ein Verkäufer eine Qualität von 5 und wünscht einen Preis von 35. Sein Käufer bezahlt einen Preis von 30.

Das Einkommen des Käufers ist:
Das Einkommen des Verkäufers ist:

Aufgabe 3:
In Runde 8 wählt ein Käufer einen Preis von 9 und wünscht eine Qualität von 4. Sein Verkäufer wählt eine Qualität von 1.

Das Einkommen des Käufers ist:
Das Einkommen des Verkäufers ist:

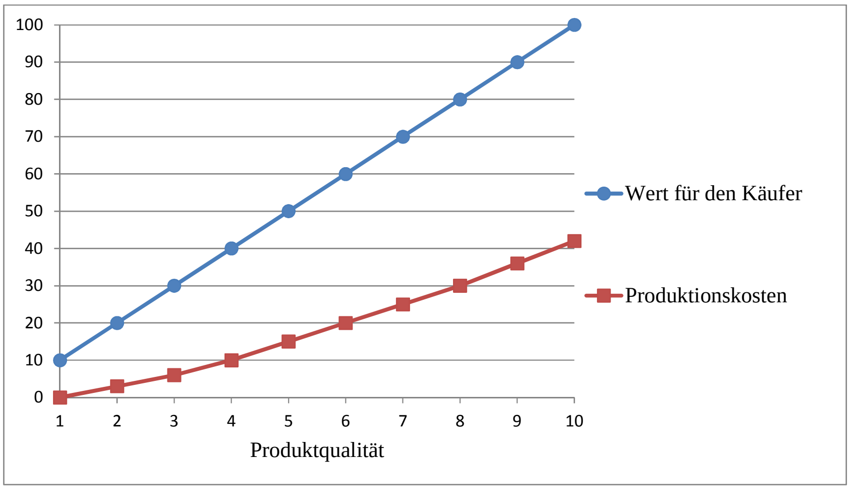
Aufgabe 4:
In Runde 11 wählt ein Verkäufer eine Qualität von 10 und wünscht einen Preis von 70. Sein Käufer bezahlt einen Preis von 75.

Das Einkommen des Käufers ist:
Das Einkommen des Verkäufers ist:

Bitte geben Sie den Studienleitern per Handzeichen Bescheid, wenn Sie mit dem Lösen der Übungsaufgaben fertig sind.
Bis zum Beginn der Studie empfehlen wir Ihnen nochmals zu überlegen, welche Entscheidungen Sie in der Studie treffen wollen.

Zusatzblatt

Qualität	1	2	3	4	5	6	7	8	9	10
Produktionskosten	0	3	6	10	15	20	25	30	36	42
Wert für den Käufer	10	20	30	40	50	60	70	80	90	100



8